Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

Potential Output, Output Gap and Great Recession in the Eurozone

Author: **Bc. Ladislav Habiňák**

Supervisor: **PhDr. Jaromír Baxa, Ph.D.**

Academic Year: **2015/2016**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, July 20, 2015

Signature

Acknowledgments

The author is grateful especially to his supervisor PhDr. Jaromír Baxa, Phd. for providing useful advice consistently and with patience. Moreover, immense support was obtained from the author's family and his Renáta for which he would like to express much gratitude.

Abstract

The thesis regarding output gap estimation is divided into two sections. The first part evaluates the latent potential output of euro area since 1998. The emphasis is put on the inquiry of usefulness of such estimates. The main findings resulting from this analysis are that, while ex-post assessment of potential output can serve as an effective tool for description of economy's behaviour in the past, the estimates evaluated in real-time are surrounded by huge amount of uncertainty which causes them to be of low reliability. For example, when searching for a structural break in the development of potential output, estimates of all models lay in one year range suggesting it to happen approximately at the end of 2007. On the other hand, the directions of output gaps evaluated at the end of real time data vintages were the same only 60% times. The second part of the thesis concerns with applicability of output gap estimates to inflation forecasting. The results show very little or no added value of such predictive modelling as autoregressive models of inflation perform comparable or significantly better forecasts for the euro area in medium and short term.

JEL Classification	E31, E32, E37, E58
Keywords	output gap, potential output, real-time data, inflation forecast

Author's e-mail	`55169537@fsv.cuni.cz`
Supervisor's e-mail	`jaromir.baxa@fsv.cuni.cz`

Abstrakt

Práce pojednávající o mezeře ve výstupu je rozdělena do dvou sekcí. První část hodnotí nepozorovatelný potencíalní produkt Eurozóny od roku 1998. Důraz je kladen na zkoumání užitečnosti jeho odhadů. Hlavním poznatkem vycházejícím z této analýzy je, že zatím co zpětné odhady potencíalního produktu mohou sloužit jako efektivní nástroj pro popis předchozího vývoje ekonomiky, odhady uskutečnené v reálném čase jsou obklopeny velkou mírou nejistoty, která tímto výrazně poškozuje jejich spolehlivost. Například veškeré použité modely odhadly strukturální změnu ve vývoji potenciálního produktu v rámci jednoho roku, přičemž naznačují, že se tato změna odehrála přibližně na konci roku 2007. Na druhé straně, směrování mezer ve výstupu v posledních kvartálech datových vzorků odhadnutých na reálných datech se mezi použitými modely shodovalo jenom na 60% celého výberu. Druhá část teze se zabývá použitelností mezer ve výstupu k předpovědi inflace. Výsledky ukazují velmi malou, nebo žádnou přidanou hodnotu takovéhoto prediktivního modelováni, jelikož autoregresivní modely inflace prokázaly podobné, nebo významně lepší předpovědi pro Eurozónu ve středně a krátkodobém horizontu.

Klasifikace JEL	E31, E32, E37, E58
Klíčová slova	mezera ve výstupu, potenciální produkt, data v reálném čase, inflační predikce

E-mail autora	`55169537@fsv.cuni.cz`
E-mail vedoucího práce	`jaromir.baxa@fsv.cuni.cz`

Contents

List of Tables ix

List of Figures x

Acronyms xii

Thesis Proposal xiii

INDUCTION - The subject of output gap 1

1 Introduction 2

2 Literature Review 5

2.1 Output gap damage during crisis 6

2.2 Uncertainty and inflation forecasting 8

2.3 Policy implications 10

ANALYSIS: PART I - Output gaps 12

3 Potential output estimation methods 13

3.1 Introduction 13

3.2 Univariate statistical method (HP) 14

3.3 Theory based approach (PF) 15

3.4 Statistical model with structural restrictions (SVAR) 17

4 Inputs to potential output estimation 19

4.1 Real GDP 19

4.2 Labour data 23

4.3 Capital and investment 28

5 The Evolution of output gap in euro area **32**
5.1 Hodrick-Prescott filter 32
5.2 Production function 34
5.3 Structural vector autoregression 36
5.4 Ex-post assessment and structural break 38

ANALYSIS: PART II - Uncertainty and inflation forecasts **40**

6 Forecasting methodology **41**

7 Inflation data **43**

8 Dispersion decomposition of output gap estimates **46**
8.1 Uncertainty within models 46
8.2 Uncertainty between models 50

9 Inflation forecasting power of output gap **52**
9.1 Harmonized Index of Consumer Prices (HICP) index 52
9.2 GDP deflator 53

SYNOPSIS - Discussion and summary **55**

10 Drawbacks **56**

11 Conclusion **58**

Bibliography **66**

LIST OF APPENDICES **I**

A Importance of output gap and potential output **II**
A.1 Sollow-Swan model II
A.2 Okun's law IV
A.3 Taylor rule V
A.4 Phillip's curve VI
A.5 Triangle model of inflation VIII

B SVAR methodology **IX**

C Method for identification of structural break **XIII**

D Dispersion of real-time vintages in time **XV**

E Development of potential output estimates **XVI**

List of Tables

2.1 Sources of uncertainty of output gap estimates 9

5.1 Descriptive statistics of Hodrick-Prescott (HP) gaps 34
5.2 Descriptive statistics of Production Function (PF) gaps 36
5.3 Descriptive statistics of Structural Vector Autoregression (SVAR) gaps 38
5.4 Structural break identification in potential output 38
5.5 Average growth of actual and potential output 39

7.1 Summary statistics of inflation in EU 18 45

8.1 Statistics of real-time dispersion from HP gaps 48
8.2 Statistics of real-time dispersion from PF gaps 49
8.3 Statistics of real-time dispersion from SVAR gaps 49
8.4 Percentage of same output gap directions 50

9.1 Student's p-values HICP index 52
9.2 Rank test p-values HICP index 53
9.3 Root mean squared errors HICP index 53
9.4 Student's p-values GDP deflator 53
9.5 Rank test p-values GDP deflator 53
9.6 Root mean squared errors of GDP Deflator 54

List of Figures

4.1 EU 18 Real-time GDP series in constant prices 20
4.2 GDP growth of EU 18 real-time vintages 21
4.3 Utilized GDP series 21
4.4 Magnitude of pre-quarter data revisions 22
4.5 Participation rate in EU 18 23
4.6 Employment in EU 18 24
4.7 Unemployment and Non-accelerating wage rate of unemployment (NAWRU) in EU 18 25
4.8 Average weekly working hours in EU 18 26
4.9 Average weekly working hours in EU 18 27
4.10 Average weekly working hours in EU 18 28
4.11 EU 18 gross fixed capital formation 29
4.12 Share of investment on GDP 29
4.13 Net capital stock in EU 18 30

5.1 Output gaps from Hodrick-Prescott filter 33
5.2 Output gaps from Production function 35
5.3 Output gaps from Structural vector autoregression approach . . 37
5.4 Potential output in EU 18 38

7.1 Development of price indexes for EU 18 44
7.2 Development of price indexes for EU 18 44

8.1 Development of price indexes for EU 18 47

A.1 Okun's law in euro area IV
A.2 Phillips curve in euro area VII

D.1 HP filter dispersion XV
D.2 HP filter same sign evaluation XV

D.3 PF dispersion XV
D.4 PF same sign evaluation XV
D.5 SVAR dispersion XV
D.6 SVAR same sign evaluation XV

E.1 September 2006 vintage XVI
E.2 September 2007 vintage XVI
E.3 September 2008 vintage XVI
E.4 September 2009 vintage XVI
E.5 September 2010 vintage XVI
E.6 September 2011 vintage XVI
E.7 September 2012 vintage XVII
E.8 September 2013 vintage XVII
E.9 September 2014 vintage XVII

Acronyms

ADF Augmented Dickey-Fuller

BK Baxter-King

EC European Commission

GDP Gross Domestic Product

HICP Harmonized Index of Consumer Prices

HP Hodrick-Prescott

IMF International Monetary Fund

KPSS Kwiatkowski-Phillips-Schmidt-Shin

NAIRU Non-accelerating inflation rate of unemployment

NAWRU Non-accelerating wage rate of unemployment

NKPC New-Keynesian Phillips Curve

OECD Organization for Economic Co-operation and Development

OLS Ordinary Least Squares

PF Production Function

RMSE Root mean squared error

SVAR Structural Vector Autoregression

TFP Total Factor Productivity

VAR Vector Autoregression

Master's Thesis Proposal

Author	Bc. Ladislav Habiňák
Supervisor	PhDr. Jaromír Baxa, Ph.D.
Proposed topic	Potential Output, Output Gap and Great Recession in the Eurozone

Motivation The dynamics and development of output gap in the context of Great Recession is of great importance. The potential success of stabilization policies largely depends on its correct evaluation. It is also vital to re-estimate the gap between real and potential output in time in order to inquire possibilities of several risks, such as inflation risk. Also, it is vital for policymakers to understand whether output and productivity loss during a recession is temporary or permanent.

Fernald (2012) suggests that at least half of the decline in actual output was caused by a reduction in its potential during the Great Recession in USA. Our motivation for this work is to examine whether similar effects occurred in Eurozone. Marcellino and Musso (2010), based on several pre crisis euro area output gap estimates, found evidence of usefulness of those estimates for real GDP growth forecasts, particularly in short term. On the other hand, they argue that there is only weak evidence on the link between output gap estimates and inflation. We will estimate the size of the gap in time to find whether our results support this evidence as one of the possible contributions to low inflationary pressures might actually be huge output gap.

Hypotheses

Hypothesis #1: The economy of euro area is still well below its potential.

Hypothesis #2: A structural break is present in the potential output at the time of the Great Depression.

Hypothesis #3: Real-time inflation forecasting using output gap estimates does not significantly enhance the performance of autoregressive models.

Methodology The first step of our research will be collection of papers concerning measurement of potential output. The main objective of this part of thesis will be to identify the best possible model or models which will be employed in empirical part of the work. We will review the literature from two perspectives. Firstly, we will consider technical features of the model selection. As one of our goals is to forecast the potential output for the coming years, we will focus on models based on economic theory. The emphasis of this analysis will be put on advantages and disadvantages of the respective models in the perspective of available data, labor market features, heterogeneity of euro area regions etc. 'Mechanical' models or filters will also be applied, however, as findings from these models suffer from end sample biases, their future potential output predictions will not be of interest. On the other hand, those models can have strong descriptive power within the sample and help identify possible structural changes. Secondly, we will review existing papers on potential output measurement of euro area. We will review the frequency of models used. We will also save the results of potential output measurements which will be compared to our findings later.

After choosing the most suitable framework, we will build empirical model mostly on OECD data. The first hypothesis will be tested by comparing our model results to the real output. The second hypothesis suggests that there was a permanent effect of The Great Recession on the potential output. Chow test will be applied in order to determine structural break in the time series. To test the third hypothesis, the methodology used by Marcellino and Musso (2010) will be applied.

Expected Contribution Given the fact that output gap and potential output are latent variables, the discussion regarding their "correct" estimation will always be present. While the work will provide the most current estimates of output gap, similarly to Anderton et al. (2014), the main added value of its qualitative assessment will lie in determination of structural break date around the Great Depression. This was rather more difficult in the past as the time series was not yet long enough to provide statistical significance of the tests.

Another area of output gap estimation which is widely analysed is regarding its uncertainty when estimated in real-time. Most of the research in this field implicitly connects this uncertainty directly to inflation forecasting such as Marcellino and Musso (2010) or Proietti et al. (2007), however, this work will focus directly on the analysis of dispersion between real-time and ex-post estimates.

The last part of results will provide euro area's inflation forecast when incorporating output gap into autoregressive models. The methodology will be closely connected to Marcellino and Musso (2010), who performed this analysis on the case of euro area as well. Nevertheless, the thesis will re-estimate the models using after-crisis data which may change the findings of these authors suggesting no added value of output gap when predicting inflation in real-time.

Outline

1. Introduction: short induction into the topic.
2. Literature Review: summary of the previous research on the studied topic.
3. Methodology: explanation of the models used in the work.
4. Data: description of the inputs into models regarding potential output estimation and inflation prediction.
5. Results: discussion of the findings resulting from models of output gap estimation and assessment.
6. Conclusion: synopsis of findings and their implications for policy and future research.

Core bibliography

Anderton, R., Aranki, T., Dieppe, A., Elding, C., Haroutunian, S., Jacquinot, P., Jarvis, V., Labhard, V., Rusinova, D. and Szorfi, B., 2014. Potential Output from a Euro Area Perspective. ECB Ocasional paper series, **No. 156 (November 2014)**.

Cerra, V. And Saxena, Sweta C., 2000, Alternative Methods of Estimating Potential Output and the Output Gap: An Application to Sweden. International Monetary Fund, **WP/00/59**.

Fernald, J., 1999. Productivity and Potential Output before, during, and after the Great Recession. Federal Reserve Bank of San Francisco, WP **2012-18**.

Giorno, C., Richardson, P., Deborah R. and van den Noord, P., 1995. Potential Output, Output Gaps and Structural Budget Balances. OECD Economic Studies, **No. 24, 1995/1**, pp. 168-209

O'Mahony, M. And Timmer, Marcel P., 2009. Output, Input and Productivity Measures at the Industry Level: The EU KLEMS Database. The Economic Journal, **119 (June)**, F374-F403.

Marcellino M., Musso A., 2010. Real Time Estimates of Euro Area Output Gap: Reliability and Forecasting Performance. ECB WP Series, **No. 1157 (Feb 2010)**.

Proietti, T., Musso, A. and Westermann, T., 2007. Estimating Potential Output and Output Gap for the Euro Area: A Model-Based Production Function Approach. Empirical Economics, **(2007) 33:** 85-113.

Author

Supervisor

INDUCTION

The subject of output gap

Chapter 1

Introduction

The dynamics and development of output gap as well as determination of components of potential output are of great importance. Over the course of business cycles, the potential success of stabilization policies largely depends on correct evaluation of size and sign of the deviation of the output from its potential. For example, if the actual output is above the potential, thus creating output gap, expansionary fiscal or monetary policy may lead to extensive inflation with only small effect on real economy. On the other hand, when the output gap is negative; the policy-makers can possibly boost the economy and more effectively reduce the economic hardship which is usually present during these periods. As potential output can be structured to several components by various approaches, the policy-makers can target those sectors of economy which are relatively under-performing.

It is moreover vital to re-estimate the gap between actual and potential output in time in order to inquire possibilities of several risks, such as above mentioned inflation risk. In the context of the Great Recession, a question which received somewhat lower attention since 1930s arose. Given that potential output is considered as a long-term trend is it possible that a huge shock to the economy as happened in the last years of past decade could cause a structural change in the potential output so that its growth is permanently lower? Even more importantly, is there a possibility under these circumstances that an intervention by the government and central bank could affect the potential so that not only the real economy but also trend recovers? If this was true, the policy-makers would face opportunity costs each year by not doing so and in the long-term perspective, there would be huge losses to the economy caused by the hysteresis effect. To some extent, these questions were inquired by Anderton *et al.* (2014) in his work related to the output gap of euro area where significant hysteresis effects were found.

The main problem of output gap estimation is that the variable is latent i.e. unobservable while at the same time it is utilized in economic modelling techniques, such as Taylor rule and other described in Appendix A. This causes large inconsistencies especially in the real-time estimates, as shown by Orphanides & van Norden (2002). Moreover, uncertainty surrounding the estimates is distributed along several dimensions described in the text. Despite some theoretical background, the relationship between output gap and inflation has been opposed by e.g. Marcellino & Musso (2010) in terms of inflation forecasting power of real-time output gap estimates. Given the research by e.g. Cerra & Saxena (2000), another source of uncertainty may come from model selection. While the institutions such as International Monetary Fund (IMF), European Commission (EC) or Organization for Economic Co-operation and Development (OECD) use production function as a benchmark model, this work incorporates three different approaches to potential output estimation in order to capture the problem of model uncertainty as well.

The analysis conducted for the purposes of this text was therefore focused on three main areas. The possible hysteresis effects of the Great Depression were firstly inquired suggesting stagnation of euro area's potential since 2007. This is in line with results of Anderton *et al.* (2014), however, this thesis also utilized formal statistical tests to identify the date of structural break which happened between 2007 and 2008. The findings based on ex-post data were consistent across models suggesting their usefulness in the description of economy. Secondly, quantification of errors resulting from using real-time vintages was performed. While the most of previous research studies uncertainty of real-time estimates implicitly through e.g. inflation forecast such as Proietti *et al.* (2007) , Orphanides & van Norden (2002) or Marcellino & Musso (2010), this work put emphasis directly on the dispersion of output gap results. The results show that all models agree on the direction of output gap in the last quarter of a vintage only approximately 60% times and, moreover, the end-sample biases of the real-time vintages within models are not seldom biased by over 1 percentage points. This results in confirmation of unreliability of real-time estimates. Finally, the thesis also evaluated inflation forecasting power of output gap resulting in no added value of incorporating output gap variable into the autoregressive model, consistently with findings by Marcellino & Musso (2010).

The work is structured as follows. Chapter 2 reviews existing literature on estimation of potential output, output gap and its inflation forecasting power. The thesis is then divided into two parts where first one captures the topic of output gap estimation and the second discusses the reliability of these estimates. Therefore, Chapter 3 and Chapter 6 capture the methodology necessary for performing the respective analyses, Chapter 4 and Chapter 7 provide description of data and Chapter 5, Chapter 8 and Chapter 9 display their results. Finally, Chapter 10 discusses the drawbacks and possible extensions of the thesis and Chapter 11 draws conclusions from results of previous analysis.

Chapter 2

Literature Review

Typically, potential output is considered as the level of output which is consistent with stable inflation and associated with other structural and institutional factors, IMF (2009b). The most prominent methods, used in official statistics at national and international institutions to estimate the potential output are the ones incorporating economic theory such as the production function approach. Alternatively, as stated by Kiley (2010), it can also be defined as a long-run stochastic trend in the time series of Gross Domestic Product (GDP). This can be obtained using statistical methods such as filters (e.g. Hodrick-Prescott filter, Baxter-King filter) which separate the trend and cyclical component of the series. Output gap is a variable which is obtained directly from estimation of potential output. It is simply the difference between the actual value of GDP and its potential

In this chapter, the review of literature is provided regarding the application and issues of output gap estimates. Firstly, the findings of prominent authors regarding hysteresis effects are summarized in Section 2.1. Then, the development of academic research regarding output gap's inflation forecasting power in time is provided in Section 2.2. Closely related to this problem, the results from literature regarding the uncertainty of estimates as well as their real-time properties are presented in the same section. Finally, the most important part of literature review can be found in Section 2.3, where policy implications from previous findings are discussed.

2.1 Output gap damage during crisis

The most recent study of development of euro area potential output is the work by Anderton *et al.* (2014) which also assesses the impact of the crisis. They conclude that crisis damaged all components of PF, however, Total Factor Productivity (TFP) was affected relatively less than labour and capital which suffered mainly by unfavourable demographic development, raised structural unemployment and declined investment rates. In the medium-term, in order for these temporary shocks not to become permanent, implementation of structural reforms is suggested by the authors. From the long-term perspective, to elevate the sustainable growth rate, they conclude that more structural reforms need to be designed and undertaken in order to offset the effect population aging. The work moreover compares different options of estimating the potential and therefore its comprehensive summary of methods is used as a reference frequently throughout methodological part of this thesis.

Proietti *et al.* (2007) provided pre-crisis analysis of potential output and output gap in the euro area by performing unobserved component models estimation. They found that estimates of output gap lay in a close range using various approaches. Moreover, the hysteresis effect was present in the labor markets. As mentioned earlier, one of the findings was that output gap is a valuable inflation predictor.

Another recent paper discussing the long-term damage of the crisis is the work by Ball (2014), where the hysteresis effects were inquired in 23 OECD countries. These effects were measured by the comparison of institutions' potential output path estimates from 2007 and today. He concludes that there is a strong hysteresis effect present in most of the countries as the potential output and actual output fell from the pre-crisis trend by almost one to one. Moreover, in the worst cases, the potential growth is depressed; therefore the losses grow over time.

Considering the mechanisms through which potential output is reduced during recessions, this issue was inquired e.g. by IMF (2009b), Reifschneider *et al.* (2013) and Hall (2014), the last two regarding the U.S. economy. The work by IMF (2009b) analyses the effects of pre and post-crisis conditions on output losses of 88 banking crisis from 1970s to 2002.The authors regress several lagged variables on the output as a percentage of pre-crisis trend. Among the most significant pre-crisis determinants according to small-scale regressions these variables were identified: share of investment on GDP, share of current account on GDP, inflation gap, fiscal balance gap, the fact whether a currency crisis emerged at the same time, share of financial openness on GDP, pre-crisis output and first year output change[1]. The significant post-crisis determinants were real government consumption growth, change in capital liberalization, change in financial liberalization, external demand shock, pre-crisis output and first-year output change. The work additionally concludes that medium-term output losses may be mitigated by proactive domestic macroeconomic policies.

Tackling the problem from different perspective, Hall (2014), decomposed the 13 percentage point decline below long-term trend in the U.S. output from 2013 into the factors of production function. Labour force participation fell by 2.4 percentage points through discouragement of individuals, growing Social-Security disability benefits dependence and higher earnings of primary family earners. The second smallest contribution, by 3 percentage points, is attributable to labour market lingering slackness; there is evidence of unusual unemployment and below standard weekly working hours. Total factor productivity damage added 3.5 percentage points while the largest share on the fall below trend is caused by a 3.9 percentage points drop in business capital.

Concerning the TFP growth before the crisis, the previous paper is consistent with results of Fernald (2014), stating that it was a result of unsustainable boom of production and IT and therefore the slowdown after crisis was caused partly by these factors. Additionally, Hall (2014) assumes that TFP evolves as a random walk with trend meaning that shocks to this variable are highly persistent and the shortfall by 3.5 percentage points is therefore permanent or nearly permanent.

Reifschneider *et al.* (2013) suggested that the potential output of U.S. economy dropped by 7 percentage points from 2007. Moreover, they argue that hysteresis effects are asymmetric and non-reversible. A strong recovery of output growth then cannot offset the labour force participation decline and unemployment jump which occur in the recession. On the other hand, other economists e.g. Ball (2014) suggest that potential output could, perhaps by policy-makers' stimulus, be raised to its pre-crisis trajectory as investment, which is considered pro-cyclical, may increase capital stock which in turn creates opportunities for job seekers and also participation rates which again stimulates investment.

[1]The term 'gap' denotes the deviation of the variable from the pre-crisis historical average (years t $= -10$ to t $= -3$, where t $= 0$ denotes the crisis year) during the last three years preceding the crisis.

2.2 Uncertainty and inflation forecasting

There are many studies regarding uncertainty of the output gap estimates. Moreover, uncertainty of these estimates has multiple dimensions. According to works of Orphanides & van Norden (2002) and Camba-Mendez & Rodriguez-Palenzuela (2001), 5 sources of uncertainty can be identified. These are data revision, model uncertainty, parameter uncertainty, final estimation error and statistical revision.Orphanides & van Norden (2002) provide a detailed study comparing the output gap estimates using real-time and ex-post data. They conclude that the estimates obtained from real-time data are of low reliability. The magnitude of measurement error is moreover compounded by the fact that data revisions are highly persistent. In contrast with Proietti *et al.* (2007), the authors found that multivariate unobserved component models generally do not provide any improvement compared to univariate models, when estimated in real time.

Moreover, according to Orphanides & van Norden (2002), unobserved components models on one hand incorporate more information sources, on the other hand, additional parameter instability and uncertainty is introduced. The biggest problem, as concluded in the study, is the unreliability of output trend estimation at the end of sample, therefore; even if the real-time data improved in quality, the end-of-sample estimates would remain of high degree of uncertainty. Finally, the authors call for great caution in the usage of output gap estimates as when these are correct, the policy-makers can greatly stabilize the fluctuations in the economy, however, when estimated imprecisely, the same actions may cause huge instability.

Camba-Mendez & Rodriguez-Palenzuela (2001) assessed the reliability of different measures of output gap for (at the time) Euro-11 area and U.S. using unobserved component model, Vector Autoregression (VAR) and SVAR approaches. They checked the consistency of sequential together with final estimates and the inflation predicting power of the models. The sequential estimates were based on quasi-real time data which are defined as real-time estimation using ex-post data. They conclude that while their multivariate specification, consistent with cyclical indicators such as capacity utilization, show temporal consistency between final and sequential estimates, this model's inflation forecasting ability is beaten by arbitrary VAR model.

On the other hand, their unobserved components model satisfied the minimum requirement; to beat random walk with trend in inflation forecasting power and although not beating univariate models with appropriate amount of lags, the authors conclude that unobserved components is a superior model to VAR specifications as these do not provide interpretable output gap measures. Although this work only deals with statistical shortcomings set by Orphanides & van Norden (2002) as it uses quasi-real time data and the dimension of uncertainty connected to data revisions is left unsettled, however, as given by Camba-Mendez & Rodriguez-Palenzuela (2001), this issue is only of relatively smaller significance.

Finally, Marcellino & Musso (2010) find that real-time output gap estimates are not particularly useful for inflation forecasting as their inclusion into autoregressive models does not bring significant improvement of the inflation prediction. Their findings are based on the euro area dataset which is convenient for comparison with findings in this work.

	HP filter	Production function	SVAR
Data revision	Yes	Yes	Yes
Model uncertainty	Yes	Yes	Yes
Parameter uncertainty	Lambda	Implicitly	No. of lags
Final est. error	Yes	Yes	Yes
Statistical revision	Highly	Little	Little

Table 2.1: Sources of uncertainty of output gap estimates

The previous table summarizes the above mentioned uncertainties in the models used in this thesis based on the review of literature. Firstly, it is obvious that all of the models have some estimation error. The possibility of choosing the wrong model is also clearly present in all of the above cases.

Additionally, uncertainty surrounding parameter space is present regardless of the approach to the estimation. In HP filter, the smoothing parameter lambda is chosen by the researcher, although some suggestions exist for the value of lambda in different frequencies. The production function does not explicitly include any parameters; a lot depends on how the researcher chooses to determine the natural level of unemployment. Should this be done by filtering methods, then the uncertainty from HP filter implicitly moves to the production function outcomes. It should be nevertheless mentioned that because of the fact that production function model employs multiple inputs, the uncertainty caused by the filtering of unemployment series causes the final results to be less affected than by directly filtering the output. In the SVAR models, one needs to choose the number of lags. On one hand, there are some criteria, such as Akaike or Schwarz criterion, available, on the other hand, the researchers usually needs to bear parsimony rules in mind which eventually makes the decision to be a more difficult one.

2.3 Policy implications

Policy implications of output gap estimates during the Great Recession are given e.g. by Bouis *et al.* (2012) who suggest that as the estimates consist mostly of TFP gaps during the crisis, monetary policy should use alternative inflation pressure indicators more heavily. Among these indicators are trends in unit labour costs, wage settlements and inflation expectations indicators. The authors conclude that until inflation expectations remain anchored, the deflationary spiral should not be triggered despite large unemployment gap in the United States. Regarding fiscal policy, they still consider output gap as a necessary input, as mechanisms through which the policy operates are robust to output gap uncertainty.

A paper by Frank Smets (1998) discusses the implications of the estimates' uncertainty in connection to Taylor rule. He concludes that although optimal behaviour of central banks is not affected by the output gap uncertainty in linear-quadratic framework, significant effects may be encountered in the class of restricted instrument rules e.g. Taylor rule which was discussed in 2.2. When the error in measurement is especially large, the efficient output gap in the Taylor rule falls to zero. The author also discusses the usage of simple rules such as Taylor rule by central banks. He assumes presence of favourable information flow effect so that it is very simple to communicate the actions of a central bank to public through a simple framework like this. Moreover, by increasing transparency of policy, the increased credibility may help to reduce volatility on the financial markets.

According to the existing literature, the relationship of output gap with inflation is a much more controversial than its relationship with unemployment. Following the Okun's law or work by Stock & Watson (1999), the business cycles should to a large extent be correlated to the behaviour of labour markets. Following this result, the author of this work included potential output estimation incorporating unemployment variables into the thesis. The case of flattening the Phillips curve during the 2000s' implies that inflation is somehow living its own life, perhaps even following a random walk, as stated by Atkeson & Ohanian (2001) or Stock & Watson (1999). Obviously, these should cause the lower forecasting power of the output gap, the proposition which was proven by Marcellino & Musso (2010) on the euro area data.

On the other hand, as Oinonen & Paloviita (2014) stated, euro area's the Phillips curve may have again steepened during the recovery. By re-estimating the gaps and their forecasting power in the similar fashion as Marcellino & Musso (2010), this thesis may bring the new findings regarding the behaviour of output gap - inflation relationship after the Great Depression. The results of this analysis should have certain policy implications.

As proposed by Bouis *et al.* (2012), the monetary policy should not depend on the output gap estimates after the crisis. The results of this work may show that the significance of these estimates in the light of inflation forecast have changed causing output gap to be an important tool for policy-makers or it may reinforce the already existing rich literature, led by Orphanides & van Norden (2002), concerning the uncertainty around output gap. In any case, the thesis should bring interesting results especially when focusing the attention on the change in the behaviour of the models and their forecasting power after the Crisis, which is supported by literature concerning the hysteresis effects of the Great Depression in the USA, e.g. Reifschneider *et al.* (2013) or Hall (2014).

ANALYSIS: PART I

Output Gaps

Chapter 3

Potential output estimation methods

3.1 Introduction

Generally, there are two main directions. Firstly, there are statistical techniques such as filtering which decompose time series into trend and cyclical components. The advantage of such methods is their relatively simple implementation; there are multiple software packages which give results of e.g. HP in a few seconds. On the other hand, these methods basically just filter out some frequencies from the data and therefore are not able to catch any structural changes within the sample. Moreover, they are highly dependent on the researcher's decision about input parameters (such as 'lambda', the smoother, in HP filter) and also suffer from end sample biases.

Secondly there are models based on economic theory. The one which occurs mostly in the research is production function approach used for example by OECD, IMF or EC. It considers the potential as an output where factors of production are fully employed. The economic background can be considered as bringing some 'structure' to the estimation as the estimates are reinforced by the theory. However, there is additional uncertainty in these models; total factor productivity and Non-accelerating inflation rate of unemployment (NAIRU) or alternatively NAWRU components of the function are themselves unobservable and often obtained by statistical filtering which puts this approach under some criticism for only redirecting the uncertainty to the sub-steps.

There is also possibility to combine the two approaches which results in e.g. SVAR or state-space models. SVAR incorporates statistical technique of vector autoregression and economic reasoning in determination of the restriction matrix. Unobserved components models together with Kalman filter use statistical methods for obtaining the trend and cycle while they can incorporate economic reasoning within their specifications.

3.2 Univariate statistical method (HP)

There are two most commonly used filters. First is HP filter, based on the work by Hodrick & Prescott (1997) and the second is Baxter-King (BK) filter, settled by Baxter & King (1999). Anderton *et al.* (2014) provide simple explanation of how these filters work; while HP filter minimizes the squared deviations of actual output from trend and thus separating the trend component in the time domain, BK filter, in the frequency domain, separates long-term low frequency fluctuations from short term high frequency fluctuations.

The main drawbacks of these methods are discussed by Anderton *et al.* (2014). Firstly, using the filtering to estimate trend implicitly creates assumption about the trend's (HP) or lower frequencies' (BK) existence. By this, there is a possibility of a mistake in identification the correct cycle as the filter may not choose the actual one. Secondly, these methods are highly dependent on the choice of parameters which is made directly by researcher. The choice is arbitrary to some degree as there exist guidelines on how to proceed. Third major drawback comes from the fact that the univariate methods suffer from very large end sample biases. Filtering is basically non-parametric method and as so it has poor forecasting reliability.

The Hodrick-Prescott filter utilized in this work is obtained by minimizing the Equation 3.1 by the trend term in time.

$$\underset{\tau}{\operatorname{argmin}} \left(\sum_{t=1}^{T} (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} \left[(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1}) \right]^2 \right) \tag{3.1}$$

Lambda is the weighting parameter which controls the smoothness of the trend line. High values of lambda reduce the sensitivity of a filter to short-run fluctuations and in the limit it should converge to the mean growth of the output during the specified period. On the other hand, zero lambda results in the perfect fit of trend to the actual values of the series. As suggested by Hodrick & Prescott (1997), the value of lambda for the quarterly data should equal to 1600. This specification, as shown e.g. by Kaiser & Maravall (1999), sets the length of one cycle to be approximately 8 years.

3.3 Theory based approach (PF)

Functional form can be of different types. According to Anderton *et al.* (2014) the most frequently occurring specification is that of Cobb-Douglas or constant elasticity of substitution function alternatively. Both usually include three factors of production. First factor is capital which is considered as fully employed at all times in the economy. Second is labour, measured most often in working hours or employment in persons; the former is a better measure which corrects for long-term secular decline in working hours per employee while the latter is easier to measure. The last term, total factor productivity, is the factor describing the level of technology in the economy and joint productivity of labour and capital.

The view given by this approach is structural as it is based on a supply side model well known in economic theory. This can help to identify the underlying contributions of respective factors as well as explain the forces underlying developments of growth in the medium term. As mentioned earlier, the production function approach has the main drawback of shifting the problems of univariate filtering methods to its sub-components such as NAIRU and total factor productivity.

Cerra & Saxena (2000), in their work comparing alternative approaches to output gap estimation, concluded that if there is enough confidence in the potential levels of inputs, then the production function approach would be a preferred method as the production factors would be constrained thus determining the potential output. In their work, however, they perform the analysis of alternative approaches on the case of Sweden in the period before millennium when there was high degree of uncertainty about the natural levels of unemployment. Therefore, they suggest unobserved components method as an optimal approach given that it estimates the NAIRU simultaneously with the potential output while it allows exogenous factors to affect inflation.

The methodology used in this work is adopted from the basic production function model, as given e.g. by Giorno *et al.* (1995) which was at the time utilized by OECD in their output gap estimation procedure. This model was firstly described by Torres *et al.* (1989) and Torres & Martin (1989). Since then, the production function methodology has become increasingly more sophisticated, current version used by EC can be found in Havik *et al.* (2014). The thesis nevertheless implements the simple version as the goal of this work is not to present the most complex model, but its partial objective is rather to compare the nature of different models and the basic production function can serve as theoretical based model just as well as any of its extensions. The estimation utilizes a logged version of Cobb-Douglas function as given in Equation 3.2.

$$y = \alpha n + (1 - \alpha)k + e, \tag{3.2}$$

where

y = natural logarithm of business-sector value added
n = natural logarithm of business-sector labor input
k = natural logarithm of business-sector capital stock
a = average labour share parameter
e = technological factor

All variables, except the technology can actually be observed or computed. Labour share is the share of households' income on total GDP. Labour input is a total of working hours and output is GDP. The capital stock is a more complex measure as only its every year's additional formation can be measured in the form of investment. The capital stock must therefore be computed, the author will focus on this issue in the data description. There is one unknown in this equation and therefore it can be estimated. The technology factor e, or TFP, is simply the difference between actual real GDP and its components.

After obtaining the TFP the procedure continues with evaluation of potential output. This consists of two steps. The working hours are adjusted to display their potential levels given by NAIRU (or possibly NAWRU), resulting in the potential level of labour input, $\bar{n}$. This step brings the structural component into the model, however, at a cost of additional uncertainty as NAIRU is unobserved. The other step is smoothing the TFP by Hodrick-Prescott filter, as given by Equation 3.1 thus obtaining new series $\bar{e}$. Finally, these adjusted labour and productivity terms are added back to the production function and together with already obtained capital and labour share variables, they create the potential output, $\bar{y}$, as can be seen in Equation 3.3.

$$\bar{y} = \alpha \bar{n} + (1 - \alpha)k + \bar{e}, \tag{3.3}$$

where

$\bar{y}$ = natural logarithm of potential business-sector value added

$\bar{n}$ = natural logarithm of potential business-sector labour input

k = natural logarithm of business-sector capital stock

a = average labour share parameter

$\bar{e}$ = filtered technological factor

3.4 Statistical model with structural restrictions (SVAR)

Discussion of usefulness of multivariate techniques can be found in Anderton *et al.* (2014). These methods aim to be superior to univariate filters by incorporating more information from the environment in which policy makers operate. Therefore their popularity has been rising in recent years. By using the additional time series besides GDP, such as inflation and unemployment, the multivariate approaches can derive trend by using either vector autoregression models or unobserved components methods. They can be considered as superior to univariate ones from the point that they can incorporate economic logic and theory, however, there is a drawback connected to estimation of larger sets of parameters which produces higher uncertainty. Anderton *et al.* (2014) moreover suggest caution when using these techniques in times of large unsustainable imbalances and non-inflationary output in the context of the Great Recession.

Vector autoregressive models were first introduced by Sims (1980). In these models, each endogenous variable is explained by its own lagged values and additionally by current and past values of other variables. By this, it is possible to capture a lot of dynamics in multiple time series. Cerra & Saxena (2000) in their work which summarized alternative approaches to potential output and output gap estimation include structural vector autoregression as proposed by Blanchard & Quah (1989). This is a classical specification regarding output gap and many extensions can be found in the literature e.g. King *et al.* (1991) or Bayoumi & Eichengreen (1992). The limitation of the SVAR approach is that it is able to identify only the amount of shocks as the number of variables used.

Therefore, it is often difficult to translate composite shocks to all respective economic variables.

SVAR methodology is provided in Appendix B. The description is inspired by a deep yet comprehensive summary of this method given by Claus (1999) who applied extended version of Blanchard-Quah type SVAR on New Zealand data.

Chapter 4

Inputs to potential output estimation

This chapter describes the data necessary for the construction and analysis of output gaps in this work. The author presents an in-depth inquiry into the model inputs as changes in output gap are bound to these variables. Therefore, understanding the data will help the reader realize consequences of their behaviour on potential output.

4.1 Real GDP

The analysis uses real-time GDP from OECD Original Release Data and Revisions Database[1]. As given by OECD, the database enables access to originally published values of over twenty economic variables for all OECD members, Euro area and several additional countries. The EU 18 GDP vintages of quarterly data in constant prices since September 2006 were downloaded, each starting the first quarter of 1998 resulting in 34 vintages of 34 to 67 observations. Additionally to the GDP series, the fixed capital formation in constant prices from this database was utilized in the work for the purposes of capital stock calculation. All variables were transformed into logarithmic scale

Approximately 5 observations were missing in each variable, all of them being the final observation of the vintage. They were calculated using either using extrapolation or utilization of the trend from surrounding vintages.

Although each series was published in constant prices, the base year for

[1]The database can be found through following link:
`http://stats.oecd.org/mei/default.asp?rev=1`

different vintages varied. This can be seen in Figure 4.1, where, for the sake of clarity, only every fourth vintage is shown. It is easily visible that all the series from September 2012 are in higher levels than the rest. There were totally 3 jumps in the real GDP series, some are not visible as not all vintages are shown in the graph. Because of the fact that large part our analysis is conducted on the differenced series and filtered series, the consequences of different levels, coming from the diverse base years for prices, in the series is negligible.

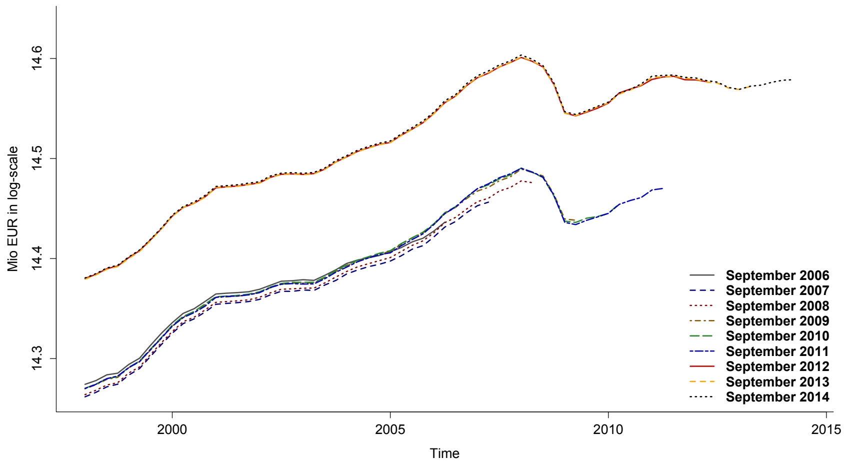

Figure 4.1: EU 18 Real-time GDP series in constant prices

The real GDP growth is captured in Figure 4.2. A period of macroeconomic stabilization from the beginning of 2000s, known also as the period of the Great Moderation, was followed by a huge crisis, where peak quarter to quarter decline in the real GDP of almost 6% occurred in June 2008. Since then, there has been a short recovery with another hit in 2011 and a slow growth afterwards.

It is clear that the GDP growth rates in the data are not affected much neither by different vintages nor by the difference base year for prices. The period of highest dispersion between vintages occurred before and during the crisis, however, the biggest change was only around half percentage point. Nevertheless, for the reason that capital formation series, needed to be set to the same base year for the purposes of calculating the capital stock, as described later, also GDP series were adjusted, as can be seen in Figure 4.3.

The adjustment process was, however, complicated because of the fact that the base year for different vintages was difficult to find. Using the deflators obtained from the OECD real-time database did not produce satisfactory results - the jumps in the original GDP data occurred at different quarters than the

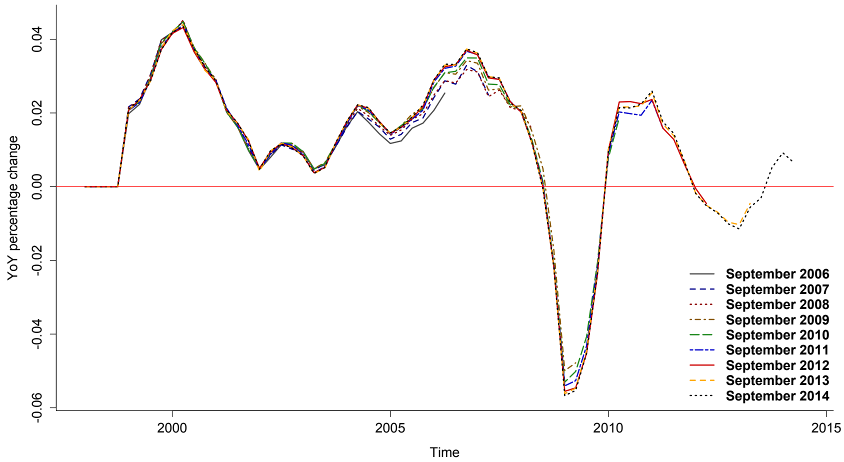

Figure 4.2: GDP growth of EU 18 real-time vintages

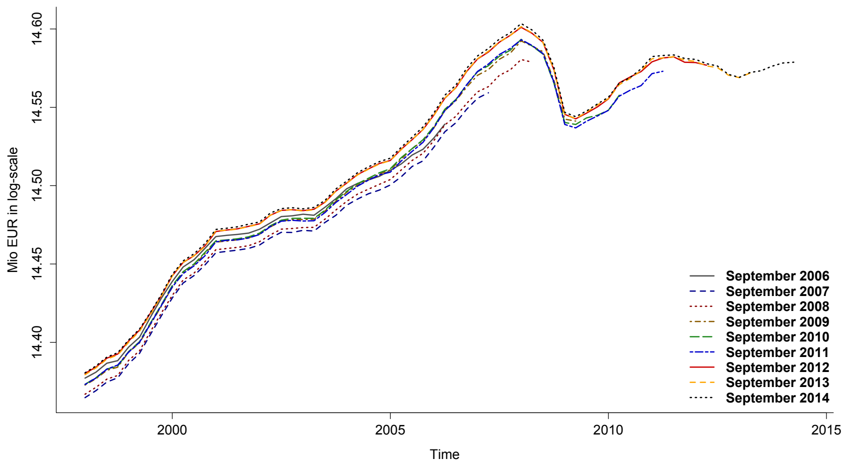

Figure 4.3: Utilized GDP series

changes in deflators' base years. Combining the information then produced even more inconsistencies between the vintages. The author therefore searched for an alternative way to rescale the series. The process for achieving the series to be on approximately the same price level was utilized as follows.

Firstly, National accounts where inquired in the OECD database where the current real GDP series in constant 2005 prices was found. This was compared to the data vintages from the real-time database. As all the vintages from the middle of the three stacks (jumps) were reasonably similar to the time series from National accounts, their base year was assumed to be 2005. Based on this assumption the other two stacks were divided by average percentage difference relatively to the middle stack to get approximately the same price level. It will be shown in the following paragraph that rescaling of the series did not change their behaviour in any way which would bias our analysis. For even further assurance, the whole analysis was also re-conducted with the original data which did not change the findings.

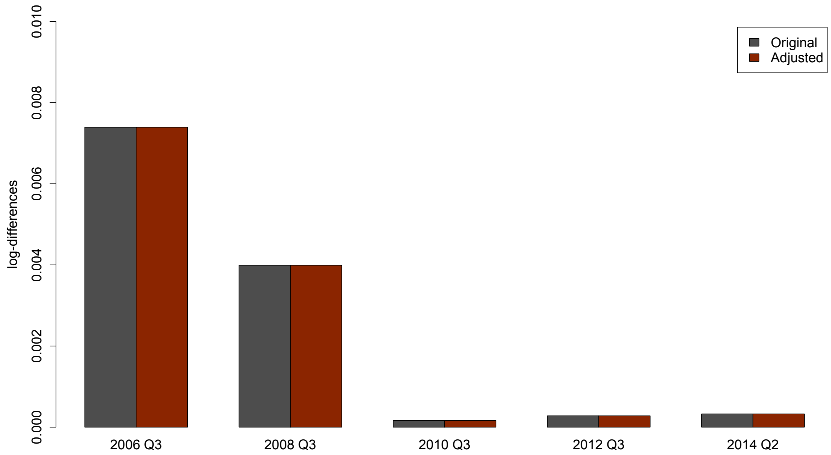

Figure 4.4: Magnitude of pre-quarter data revisions

As stated before, despite the fact the difference in levels of the original series, the adjusted series do not behave differently in terms of growth rates. Another question may be raised as to what extent the original and adjusted (utilized) series vary between different time vintages. Figure 4.4 shows the dispersion between original and adjusted series in terms of pre-quarter data revisions. For example, the graph displays that for September 2006 vintage the value of real GDP for the second quarter equalled around 0.7 percentage points higher than when it was firstly published, i.e. in June edition. This applies to both when calculating this dispersion from original as well as adjusted series. At this point, it is important to remind the reader that the scale of the y-axis is not as much important for the implications of this figure. What is important is that the utilized data, except for the years where there was a change in the base year for prices in the original series, behave the same as the original series. For the sake of brevity, the same adjustments and results mentioned above hold also for the series of gross capital formation.

4.2 Labour data

The majority of time series regarding population where downloaded from the Eurostat database[2]. The quarterly data from 1998 were used. The author inquired the EU 18 series of total population, active persons, employed persons and total working hours. Before the year 2005, however, there were several inconsistencies in the data. The major one was that first nine quarters were missing in most of the series and then there were only yearly data until 2005.

The problem was approached by summing the interpolation of all individual EU 18 countries' time series which were denominated in thousands of persons. This brought some new information to the series compared to the simple interpolation of Eurostat's EU 18 series as the data for the two biggest countries, France and Germany, were available at least yearly from Q2 1998 and additionally, data for the following two largest states, Italy and Spain were available quarterly from the beginning. Cumulatively, these four members of the euro area sum up to more than three quarters of Euro-zone's total population. For the series of working hours, the simple interpolation seemed to bring sufficiently good results compared to GDP weighted working hours of respective countries. All data had to be furthermore seasonally adjusted using the package 'seasonal' in R. Figure 4.5 shows the development of participation rate in EU 18. The series named 'EU' stands for the interpolated Eurostat series while the 'EUc' series is sum of individually interpolated series of EU 18 countries.

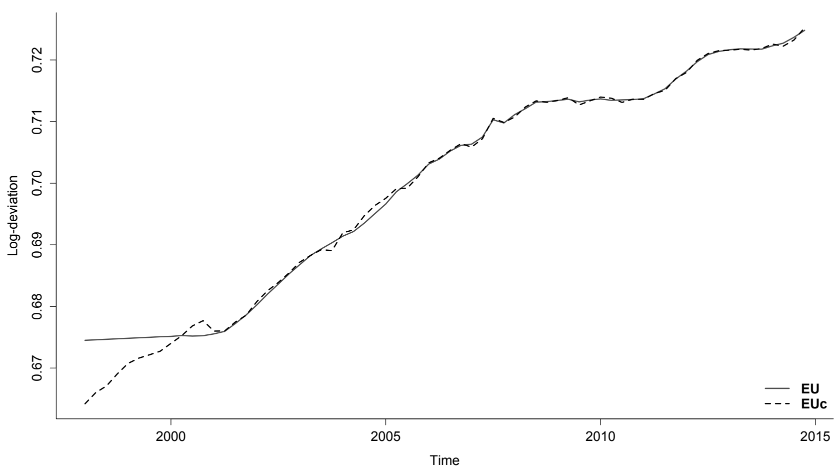

Figure 4.5: Participation rate in EU 18

[2]The database can be found through the following link:
`http://ec.europa.eu/eurostat/web/lfs/data/database`

The previous graph presents an interesting phenomenon of increasing labour participation rate in EU 18. The ratio is computed as a share of economically active persons on total population. Although there was only around four percentage points incline, it is still quite surprising, especially when put in contrast with the development of another developed economy, USA, where there has been a 3 percentage point decline in this ratio from 66 to 63 percent since 2007, as given e.g. by Aaronson *et al.* (2014). While the relative number indicates a positive trend, because of declining population size the actual number of economically active persons in the euro area oscillated around 155 million from 2008. Also, there is around one percentage bias at the beginning of the sample. Again, the recalculated data seem to be projecting a more realistic outcome as the 'EUc' series follows the long term trend.

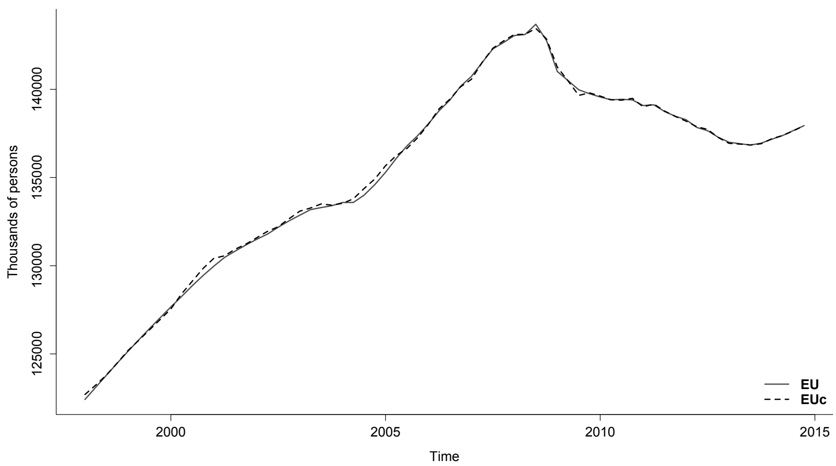

Figure 4.6: Employment in EU 18

Figure 4.6 shows the development of total employed persons in the EU 18. In this case the calculated and original series show almost exactly the same results. It can be seen that between 1998 and 2008 the number of employed people rose from around 120 to 145 million. After the crisis, however, the series suffered a decline until 2014 when it slightly grew for the first time in seven years. The underemployment lays down a lot of questions to where the potential output could be, some descriptive analysis of unemployment gap can be found in the following paragraph.

The previous chart demonstrates the evolution of unemployment in the euro area. The filtered unemployment rate exhibits a U-shape in the analyzed period. During the Great Moderation the share of people without job on active

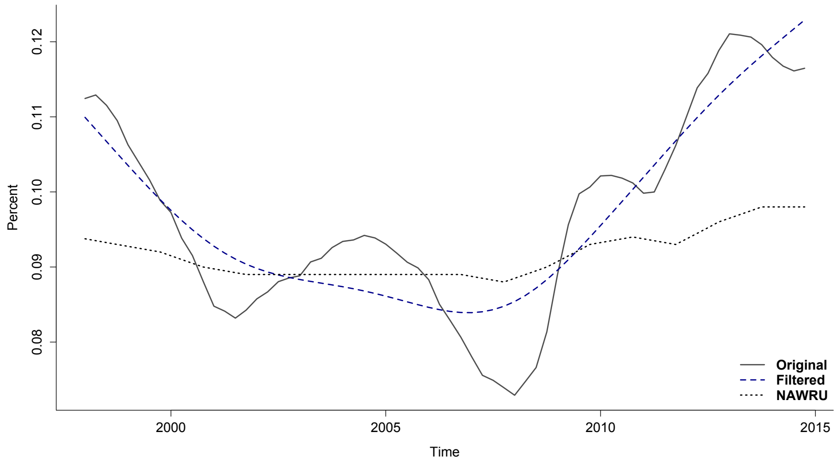

Figure 4.7: Unemployment and NAWRU in EU 18

population was falling, reaching the original series bottom value of 7.3% in the second quarter of 2008. Since then, the unemployment rose to over 12% in 2013.

Figure 4.7 also brings down the development of NAWRU. The NAWRU series was downloaded from AMECO database[3] of European Commission where the major indicators of economic activity can be found. As explained before, when actual unemployment is higher than NAWRU, the inflation pressures should appear in the economy. The non-accelerating wage rate of unemployment is supposed to be much more stable than actual unemployment. It basically serves as a labour related counterpart of what the potential output is to the real GDP. As well as potential output, the NAWRU is unobserved and there are various techniques to gain it. The author chose the estimation of EC for this series as he finds it to be a reliable source because of the information availability about euro area that the EC has. Following the previous graph, if the filtered series is taken into consideration, there were two changes in the direction of unemployment gap, defined as the difference between actual unemployment and NAWRU.

At the beginning of the analyzed period, until around 2003, the gap was positive, indicating some space to adopt measures promoting expansive policy. When analyzing the chart ex-post, there were definitely signs of overheating present in the labor market from 2003 to 2009 with the unemployment gap reaching the size of minus two percentage points in the second quarter of 2008.

[3]The database can be found through the following link:
`http://ec.europa.eu/economy_finance/ameco/user/serie/SelectSerie.cfm`

After the Great Recession, the gap became positive again and additionally also increasing. The large unemployment gap may be currently one possible source of inefficiency in the economy. This proposition comes from the way production function approach to potential output estimation is built, as mentioned in the chapter concerning methodology.

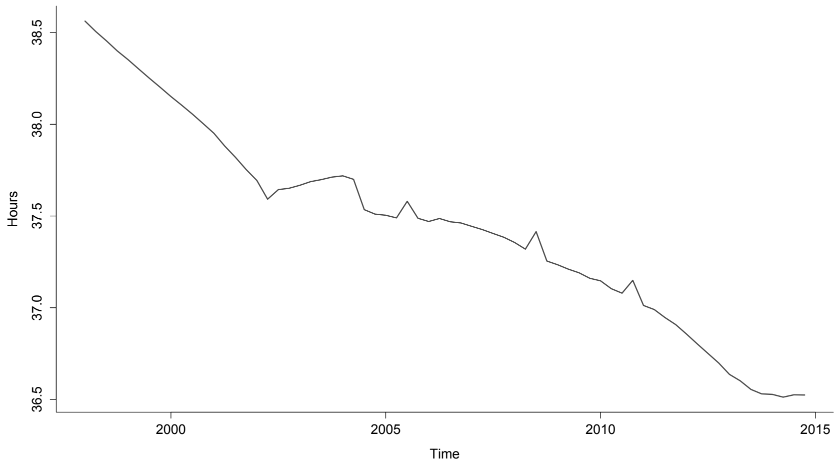

Figure 4.8: Average weekly working hours in EU 18

Figure 4.8 shows the evolution of the final indicator of labour workforce strength, the average weekly working hours in the Euro-zone. When the series was recalculated in the similar way as the series before, only using GDP country weights instead of sums, the resulting series was very similar in slope but was offset by around 2 working hours. The author therefore utilized the original interpolated series. There is a clear continuous trend of declining working hours in the economy which fell from almost 39 hours per week to 36.5 since 1998. The following chart combines all the previous findings from labour markets.

The previous graph depicts actual as well as potential average total yearly amount of hours worked in the euro area. The amount is calculated by multiplying the average working hours by total weeks in a year times employed persons or potentially employed persons, which can be obtained by utilizing the NAWRU, respectively. The resulting variables serve as an input into the production function model of this thesis and, given the labour share on GDP being to around two thirds, are therefore of large importance for the model results.

The decline in working hours seems have somewhat lower effect on the new variables than unemployment figures and projections. The unemployment data

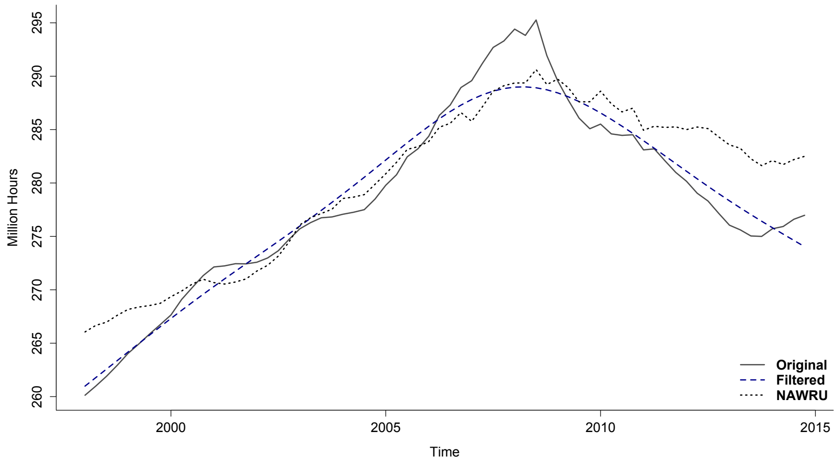

Figure 4.9: Average weekly working hours in EU 18

evolution seems to have a lot of weight assigned in computation of the final total working hours series. There is 2% difference at the beginning and 3% difference at the end of the series in filtered and potential working hours. It can be a priori assumed that these obvious divergences will cause the potential output resulting from Hodrick-Prescott filter as well as SVAR model to be undervalued relatively to the production function approach as the former two do not include potential labour in the computation.

There was one remaining parameter regarding labour which is needed for the production function approach to potential output. The parameter is a wage share, sometimes called the labour share of income, e.g. IMF (2009a). As suggested by its name, it is a ratio of employee compensation on the country's product. It is an estimate of income distribution between labour and capital and is a necessary input to production function model.

On one hand, as shown by IMF (2009a), the wage shares fell significantly in the end of the century, on the other hand, as given in the Figure 4.10, which is based on the EC data from AMECO database, the shares have been reasonably stable since the beginning of 2000s. The ratio has never crossed neither the bottom boundary of 0.6, nor the top 0.65. It fell in the first five years of the decade but rebounded to the long-term mean after the crisis. The author interpolated this series to quarterly data and utilized it in the production function estimates. For the computation of the potential output, the mean value of wage share throughout the analysed period was used.

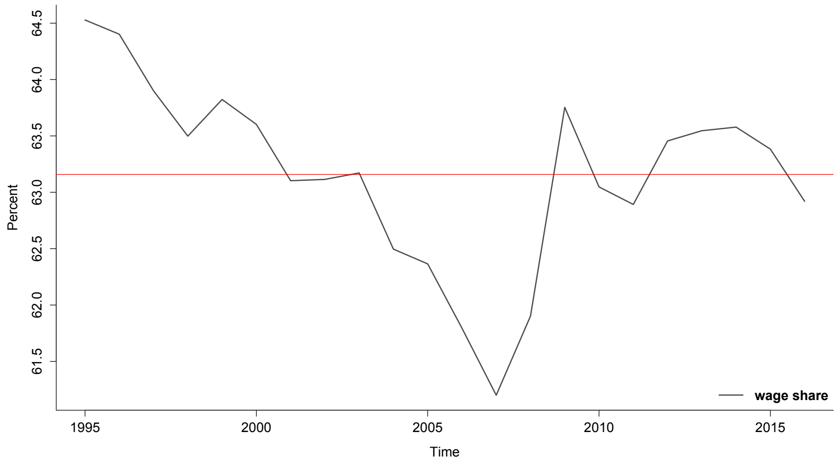

Figure 4.10: Average weekly working hours in EU 18

4.3 Capital and investment

This section analyses the behaviour of capital formation and capital stock, moreover, it describes the way the author computed the unobservable capital stock variable based on the available information and methodology.

As mentioned earlier, gross capital formation series, investment, were downloaded from OECD real-time database. The already adjusted series of gross capital formation in constant 2005 prices can be found in Figure 4.11. This graph is interesting for multiple reasons. Firstly, compared to the GDP vintages, it is visible that the data revisions concerning investment are of lower magnitude. Secondly, the figure can provide a nice description of what happened during and after the crisis. The 2008 level of investment into the economy was around 15% higher than that at the beginning of the decade. A sharp rise can be seen from 2005. When the crisis hit, however, the capital input fell significantly by almost 20% and has not recovered since then.

The next interesting finding may be that share of investment on GDP has clearly been volatile as well. The gross capital formation to GDP ratio is depicted in Figure 4.12. The mean value of this ratio equals to almost two thirds, however, the development throughout the last six years shows the drop from around 70% to approximately 58%. The stagnation of GDP be therefore, at least to some degree, be assigned to low level of investment which may be caused by e.g. unfavourable condition on financial markets because of the credit and debt crises.

It should be noted that the capital stock series is, in contrast to investment,

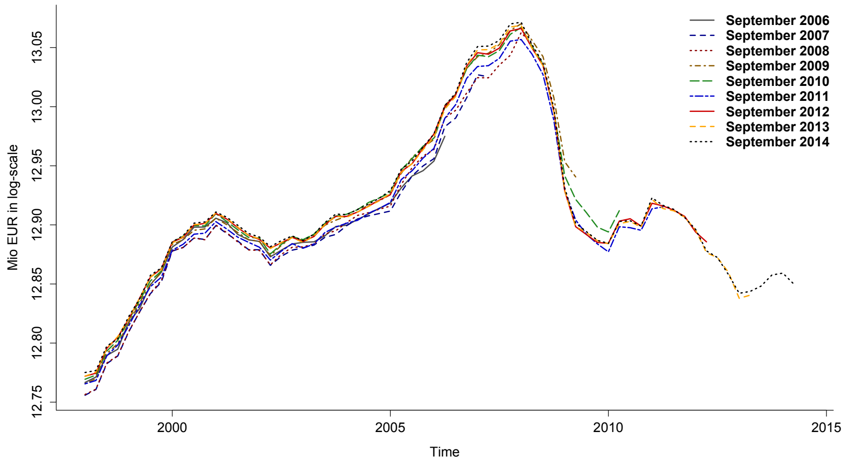

Figure 4.11: EU 18 gross fixed capital formation

much harder to obtain. There is a possibility to compute the values by a perpetual inventory method, described e.g. in OECD (2001) and OECD (2009), or use different way, such as making alternative distributional assumptions on the stock depreciation, where Winfrey distribution is commonly used, as in the paper by Derbyshire *et al.* (2010). There is also a possibility to download the already computed capital stock data from highly relevant sources. One such may be the European Central Bank database ; however, the data here are only available until 2008 which is not sufficient for our analysis. Another data source can be found in an already mentioned AMECO database. It provides the yearly net capital stock data in the required period, starting as early as 1995.

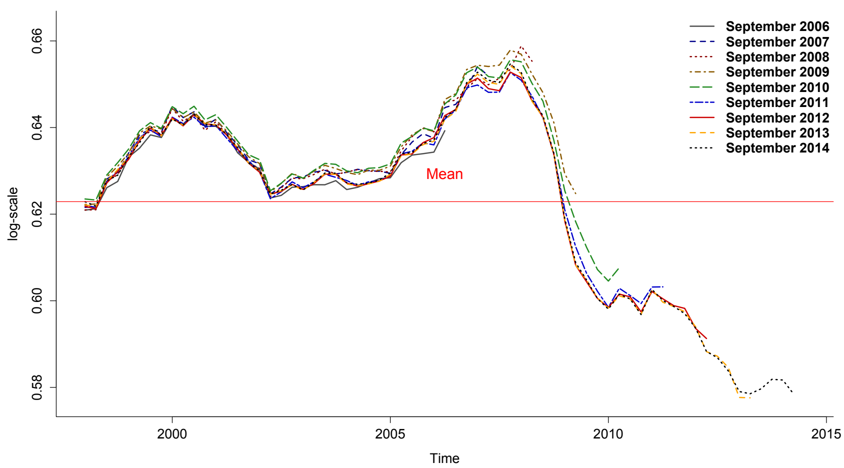

Figure 4.12: Share of investment on GDP

On one hand, if the thesis utilized those data, interpolated to quarters, some information would be lost due to the lack of short-term fluctuations which can be found in the investment data vintages. On the other hand, the EC data can serve as verification for the method finally utilized in this work. Having the way the capital stock was computed for this work explained in the methodological part, the author would now like to show the difference between his and EC's computations on the following figure.

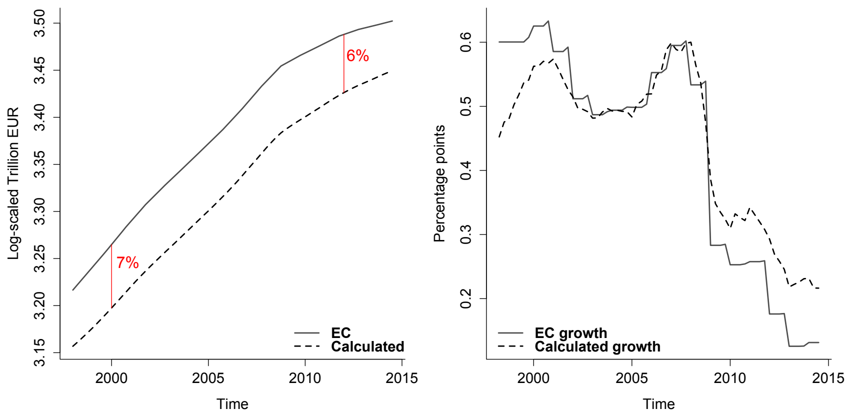

Figure 4.13: Net capital stock in EU 18

Figure 4.13 shows the development of net capital stock in the euro area as well as differences between the series downloaded from AMECO database of European Commission and the one computed for the purposes of this work. The author believes that the relative difference is fairly reasonable given the fact that the capital stock is a non-observable variable. Moreover, when considering the growth evolution of the net capital stock, as depicted on the right panel of Figure 4.13, the rates are very much similar for the majority of the observed period.

The short-term discrepancies may well be caused by the difference between interpolated EC series, represented by the solid line, and utilization of OECD real-time quarterly gross capital formation series, associated to the dashed line. Now, it is important to realize that the capital stock series is being used as an input to the production function model. The fact that the two series slightly differ in levels, rather than growth, can be considered negligible because, if one switched the two in the model, this difference would be carried over to the total factor productivity term and would be displayed simply as a shift in TFP.

The way the procedure for computing the capital stock was done is explained in the following text. Firstly, the estimates of Derbyshire *et al.* (2010) were taken as a base. These estimates were provided for all EU 18 countries for the year 1995. All applicable countries' figured were summed and transformed to 2005 prices. Then, the average capital formation between 1998 and 2000 was extrapolated to 1995. The author believes that this assumption is reasonable because as at the end of the century, constant GDP growth rates visible in Figure 4.3 and relatively stable share of investment on GDP, as can be seen in Figure 4.12, were present. The capital depreciation was the one remaining feature necessary for computing the net capital stock. It would be clearly misleading not to incorporate the capital depreciation to the model as the analysed period is as long as 17 years and one could find it hard to believe that e.g. all the machinery used in 1998 was still in production in 2014. This term was therefore calibrated according to several sources.

According to the DSGE model introduced by Adjemian *et al.* (2007), the quarterly depreciation rate for the euro area is equal to 0.0025 per quarter which is a counterpart for around 1% yearly depreciation. The more recent DSGE model, given by Quint & Rabanal (2013) suggests quarterly depreciation rate to be 0.0125, corresponding to approximately 5% depreciation per year. The results in this thesis are presented using the annual depreciation rate of 4%, although this number was experimented with and it should be noted that changing it by reasonable amount did not significantly affect the outcome of production function model. The utilized number is closer to the more recent paper which is the main reason it was chosen. Moreover, the author finds it more economically reasonable as the pre-crisis interest rates on the top rated government bonds were converging to this number and as is given in the theory, on the optimal path, investment offsets the depreciation.

Chapter 5

The Evolution of output gap in euro area

This chapter provides the reader with the results of analysis developed for the first part of the thesis. The computed potential outputs and output gaps are described for each model separately as well as compared across these methods. Special attention is given to the inquiry of structural break in the euro area GDP series.

5.1 Hodrick-Prescott filter

The following chart displays the output gaps of every fourth data vintage obtained by HP filter procedure, measured as a percentage deviation from the potential output. The gap estimates for individual vintages lie reasonably close to one another at the beginning of the analysed period. For example, the Hodrick-Prescott filter computed the output gap in the first quarter of 1998 to be minus 0.5 percent with the range of 0.05 percentage points. Similar findings can be obtained when analysing the gap results until around early mid 2000s. The end-sample bias of the first vintage in the HP filter approach to estimation is present already 13 periods before its end. While the period between 1998 and 2003 showed standard deviation of estimates between vintages to be closely around 0.1, this statistic experienced a sudden jump in the value of approximately 0.2 for observations from 2004 to 2006. Following 2006, the dispersion between the vintages increased with especially large deviation around the period of crisis. On the other hand and quite interestingly, the standard deviation of observations between vintages after 2010 reduced to around 0.15, showing smaller dispersion than in the period before 2006.

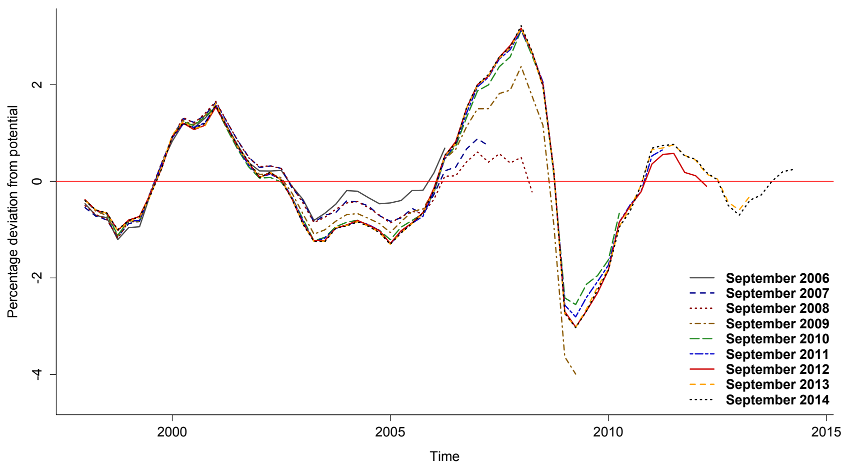

Figure 5.1: Output gaps from Hodrick-Prescott filter

The author would a priory expect similar magnitude of standard deviation at the end of the sample as during the period between 2003 and 2006 as both suffer from end-sample biases and neither one is biased by a huge one-off impact of crises. The comparison of findings from these two periods can, nevertheless, serve as a nice example of how HP filter works. The author believes that it is essential for the researcher to apprehend the mechanisms of the models she uses so that it is possible to choose the right one according to the situation in turn.

The answer to why the gap behaviour differs for the two time periods lays in the Figure 4.3, which shows the evolution of real GDP. The major difference which matters for computations is that while the first period exhibits slow growth followed by sharp rise, the second period showcases the times of recovery after which stagnation occurred. While in the former case, the sudden growth appears somehow non-expecting for the procedure after some time of smooth slow incline, which caused the gap estimates to be undervalued as they were being continuously upward, in the latter instance, the observations after the growth break were being almost constant which caused stabilization of the estimation at the end of the sample.

While the previous paragraphs were concerning mostly with the statistical results between the data vintages, the Table 5.1 summarizes some descriptive statistics within several chosen vintages. The author decided to show statistics

	Dec 2006	Dec 2008	Dec 2009	Dec 2010	Dec 2014
Min.	-1.17	-1.18	-3.37	-2.54	-3.02
Median	-0.24	-0.11	-0.26	-0.24	-0.07
Mean	0.00	0.00	0.00	0.00	-0.00
Max	1.55	1.65	2.72	3.14	3.26
Std. dev.	0.73	0.76	1.36	1.35	1.30
Obs	35	43	47	51	67

Table 5.1: Descriptive statistics of HP gaps

from the end of first and last vintage year as well as years surrounding the crisis. As can be seen, the mean value of each series was zero which comes from the set-up of the Hodrick-Prescott filter. Interestingly, the majority of displayed statistics changed after the crisis. On average, the minimum value decreased almost three times from around -1.2 to approximately -3 percent while the maximum as well as standard deviation nearly doubled; the former from over 1.5 percent and the latter from around three quarters of a percentage point. The NAs term shows how many observations were left out compared to the final December 2014 vintage which was totalling 67 quarters.

5.2 Production function

The following analysis studies the behaviour of output gaps obtained from production function approach. From the first look, the outcome of Figure 5.2 looks similar to the results from HP filter. The qualitative results which take sign of the output gap rather than its magnitude into consideration are truly similar, at least until 2011. The occurrence of negative gap at the end of the century followed by few years of overheating signs and afterwards again consecutive under-performance of the economy reminds the classical view of the business cycles repetition. The period around crisis can, from the higher perspective, be viewed as another cycle as well only with much higher peak and significantly lower bottom. The qualitative description of the two estimates so far differs after the period of recovery. While Hodrick-Prescott filter outcome suggests that the euro area is currently around its potential, the production function method reports a lot of space over the actual production.

The quantitative results, however, differ between the two approaches. The main reason for this is incorporation of NAWRU methodology into the production function estimates. Figure 4.7 shows the development of NAWRU in time

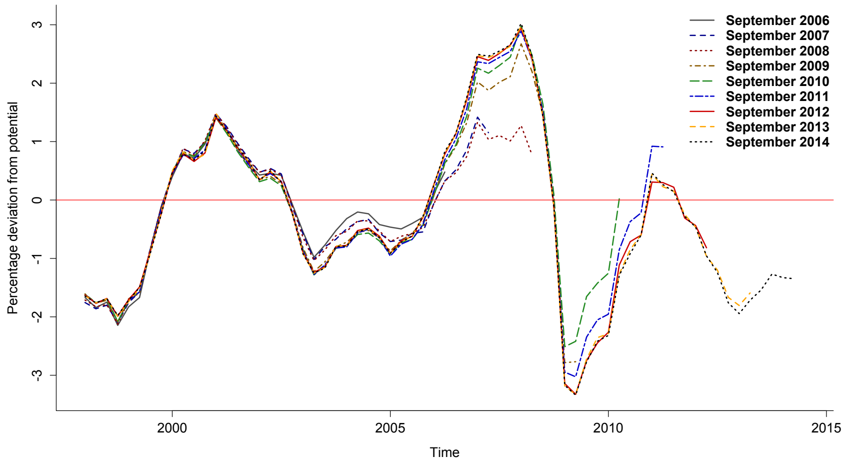

Figure 5.2: Output gaps from Production function

from which several conclusions regarding the difference between HP and PF results can be drawn. The large negative unemployment gaps at the beginning and at the end of the period pull the output gaps in the same direction. Following the same logic, the fact that NAWRU is higher than actual unemployment rate in the pre-crisis period, the output gaps are, on average higher during the time. There is one percentage point lower gap from production function in the late 90s which is actually double the amount estimated by Hodrick-Prescott filter. The following cycle from 2000 to 2006 seems to get approximately same values from the two approaches as the unemployment gap was the smallest in this period.

In addition to previous findings, the output gap vintages from production function seem to lie closer to each other during the period. The two previous charts display one partly visible result which is supported by plotting all vintages; that while HP gaps lie in the range from 0 to 3 percentage deviation from potential in the shortly pre-crisis period, the PF gaps range is smaller, starting at 1 percent and maximizing at similar numbers. Smaller variance between vintages basically means less uncertainty within the model.

During the sharp downturn, the PF gap estimates have slightly larger variance; however, there is higher number of extreme results from HP filter. The reason is that the most current vintages of HP filter estimated the gaps at the time in a very close range while the vintages from 2009 and 2010 exhibited vast amount of dispersion. The minimum value was as low as -4 percentage deviation from potential while the maximum number was 1.5 percentage points higher. The PF results, on the other hand, lie all in a range between -3.3 and -2.5 percentage deviation from potential. The situation turned around at the end of the recovery period where there were few vintages from production function which overvalued the output gap. On the other hand, their overshooting was at maximum of half percentage point, much lower than HP undershooting in the previous period. Finally, the production function suggests the negative output gap of around minus two percent these days.

	Dec 2006	Dec 2008	Dec 2009	Dec 2010	Dec 2014
Min.	-2.11	-2.13	-2.55	-2.61	-3.33
Median	-0.27	-0.01	-0.19	-0.17	-0.60
Mean	-0.24	-0.03	-0.03	-0.07	-0.34
Max	1.41	1.48	2.86	2.95	3.08
Std. dev.	0.95	1.02	1.38	1.39	1.46
Obs	35	43	47	51	67

Table 5.2: Descriptive statistics of PF gaps

The table above displays the same descriptive statistics as Table 5.1 for estimates coming from the same data vintages. The standard deviation of production function resulting series is higher than that of the HP filtered for all of the chosen vintages, although the variance of the series coming after 2008 lies in the similar level. This is caused by the higher volatility of production function gaps at the beginning of the analysed period. Additionally, because of the fact that the production model does not have constraint on the zero mean, such as HP filter, all of the series are shifted downwards on average compared to the first model. In all the displayed cases, except for the December 2009 vintage, the production function gaps have lower minimum and maximum values compared to the Hodrick-Prescott filter estimates. For the reasons mentioned in the previous paragraph, the December 2009 vintage has, following the same logic, the opposite results.

5.3 Structural vector autoregression

As was the case for the two previous models, SVAR gaps obtained from the majority of vintages suggest positive gaps at the beginning of the decade which continuously slid into negative numbers followed by a sharp rebound at around 2005. This model therefore suggests that the crisis was caused by large overheating during the few pre-crisis years. The behaviour after crisis is again qualitatively similar to HP and PF estimates, although only until 2011. One very interesting finding is that each of the three models brings a different story to where the output gap nowadays lays. While HP filter suggests approximately zero gap, the PF estimates it to be significantly negative and to increase the uncertainty, the current SVAR gap lays 2 percent over the potential GDP.

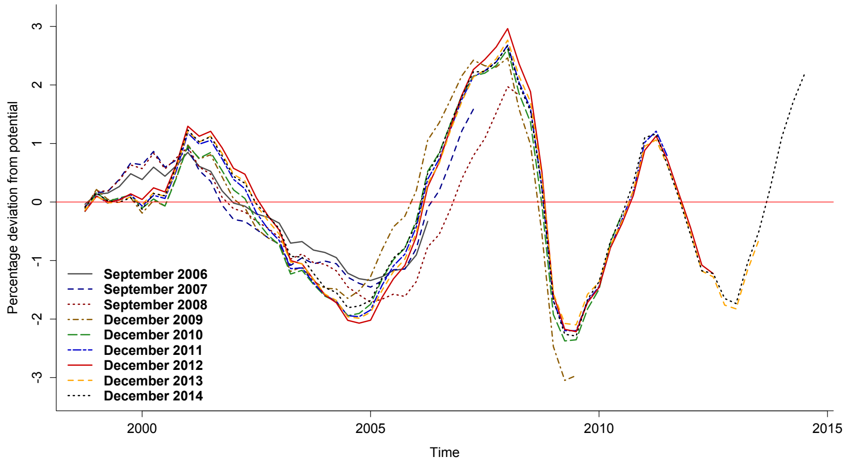

Figure 5.3: Output gaps from Structural vector autoregression approach

From the quantitative point of view, on the other hand, the situation changed quite significantly compared to the previous models. The relative closeness of individual vintages' estimates at the beginning is no more present in the SVAR output gaps. Until the crisis, these estimates were being seemingly evenly distributed in a range of around 2 percentage points. After 2010 the vintages seem to keep close to one another. The author believes that the reason lies in the number of observations used in the model. After closely examining the gaps obtained from the vintages since 2012, the dispersion between these seems to be minimal. Throughout the whole analysed period. Conclusively, the parsimony rule likely plays a large role in SVAR models. Once there are enough data points used, the SVAR real-time estimates seem to lie in the similar range as the ones resulting from the other models.

The standard deviation of SVAR within vintages estimation is a mixture of PF and HP volatility. The December 2006 series resulting statistic is smaller than for the rest of the chosen vintages, as can be seen in Table 5.3. The minimum and maximum values mostly lie inside of the range set by the other two models. Unlike the production function gaps, SVAR does not necessarily obtain negative mean and median values.

	Dec 2006	Dec 2008	Dec 2009	Dec 2010	Dec 2014
Min.	-1.35	-1.85	-3.05	-2.37	-2.29
Median	-0.10	0.00	-0.02	-0.09	0.01
Mean	-0.23	-0.03	-0.03	-0.16	-0.02
Max	0.92	2.28	2.46	2.61	2.67
Std. dev.	0.71	1.16	1.38	1.31	1.27
Obs	32	40	44	48	64

Table 5.3: Descriptive statistics of SVAR gaps

5.4 Ex-post assessment and structural break

The author presents the most important findings regarding output gap during the analysed period here. The focus will be placed on the consensual findings of all models.

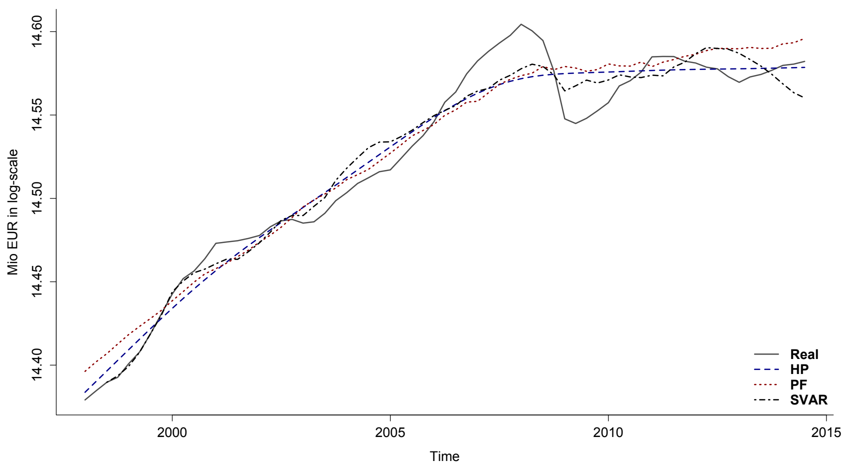

Figure 5.4: Potential output in EU 18

To assess output gap development in the euro area from the general perspective, one can come to the following findings, as given in Figure 5.4, which shows the evolution of actual and potential output. First of all, every model utilized in this work seems to suggest hysteresis effects in euro area potential output. Table 5.4 displays the results of structural break analysis.

	HP	PF	SVAR
Date	Q3 2007	Q1 2008	Q1 2007

Table 5.4: Structural break identification in potential output

While the slope coefficient in the model regarding structural break identification, as specified in Appendix C, was statistically significant at 1% level from around 2003 to 2011 for each model, all of them reached the bottom p-value in one year range between 2007 and 2008. This tightness of estimates brings additional confidence into ex-post evaluation of potential output. From the qualitative point of view, this result suggests that effects of crisis in euro area were present already in early 2008.

The following table summarizes the yearly simple growth averages computed from inter-quartile growth, as specified before, during the specified periods. It is clearly visible that the numbers differ significantly since 2008. While the average increase of actual as well as potential output was around 2% from 1998 to 2007, it has, on average, stagnated ever since.

	Real GDP	HP	PF	SVAR
98-07	2.25%	1.88%	1.77%	1.98%
08-14	-0.34%	0.10%	0.34%	-0.27%

Table 5.5: Average growth of actual and potential output

Secondly, while ex-post estimation of output gap during the crisis differs across the models, the large overheating shortly before the crisis is computed quite consistently with all methods suggesting the gap between approximately 2.5 to 3%. The final finding explored by the author is that ex-post vintages have the power to qualitatively assess the output consistently across models. All models estimated the business cycles in a similar fashion since the beginning of the century with the dispersion in the last 3 years attributable to real-time biases.

ANALYSIS: PART II

Uncertainty and Inflation Forecasts

Chapter 6

Forecasting methodology

The methododology by Marcellino and Musso (2010) was largely adopted in order to assess the inflation prediction power of the models. One quarter and one year ahead forecast horizons are considered in order to asses the power of short-term and medium-term predictions. The model is given as:

$$\pi_{t+\tau}^{(\tau)} - \pi_t = \alpha + \sum_{k=0}^{2} \beta_k \Delta\pi_{t-k} + \gamma x_t + \delta \Delta x_t + u_{t+\tau} \tag{6.1}$$

where $\pi_t^{(\tau)} = \frac{400}{\tau} \ln\left(\frac{p_t}{p_{t-\tau}}\right)$, $\pi_t^{(1)} = \pi_t$ and x_t is output gap denominated in percentage deviations from potential.

The benchmark model, nested in 6.1 is $AR(3)$ model for inflation with drift, as given by Equation 6.2.

$$\pi_{t+\tau}^{(\tau)} - \pi_t = \alpha + \sum_{k=0}^{2} \beta_k \Delta\pi_{t-k} + \epsilon_{t+\tau} \tag{6.2}$$

In contrast to Marcellino and Musso (2010), this work does not utilize the MSE-t statistics. The testing for significance was done as follows. Firstly, estimated change in inflation coming from the end of each data vintage was deducted from its actual value. Then these numbers were squared to get the series of positive numbers. The same procedure was afterwards applied to the AR3 model. AR3 model was utilized, despite the fact that the underlying paper uses AR4 simply because the fourth coefficients were always insignificant in all models. Therefore, due to parsimony rules, AR3 model was used.

Number of observations varied by the output gap estimation approach as well as by the chosen time period which was being forecasted. It was nevertheless always over 30 which, according to some researchers, is the lower bound for application of central limit theorem. Nevertheless, the author used paired one-sided t-test as well as paired one-sided Wilcoxon signed rank test with the alternative that output gap models produce smaller squared errors than autoregressive inflation forecast to obtain robust results.

Chapter 7

Inflation data

As given by the theory, the output gap should play a significant role in inflation forecasting. Large gaps are supposed to be a sign of overall overheating, causing the pressure on inflation. For this reason, the price development is a very important variable in this work. HICP as well as GDP deflator was downloaded from OECD database.

There are two main differences between these measures. Firstly, HICP index displays the price level for domestic as well as imported products while the GDP deflator only considers the domestic production. Secondly, the computation of HICP index only includes the goods and services attributable to consumers' expenditures while the GDP deflator implements all the expenditures in the economy. A priory, the deflator is considered a more relevant measure for this analysis by the author because of the two differences and also because of the fact that also the paper by Marcellino & Musso (2010), from which the methodology for testing of real-time output gaps inflation forecasting power was largely replicated, uses it as an inflation proxy. Nevertheless, this thesis shows results of inflation forecasts from using both indexes as the author believes that this may bring some interesting findings exactly because of the differences between them. The consumer prices may behave differently especially in periods of unusual macroeconomic conditions. Additionally, the prices of imported goods may, from the first thought, be immune to the domestic overheating of the economy; on the other hand, the importers more dependent on the publishing of real-time figures of e.g. production or income in the economy which may in turn cause the real-time gaps to be good predictors of the importers' behaviour.

The previous chart displays the HICP index and GDP deflator progress since the third quarter 1996. Both indices have the base year set to 2005. Unlike

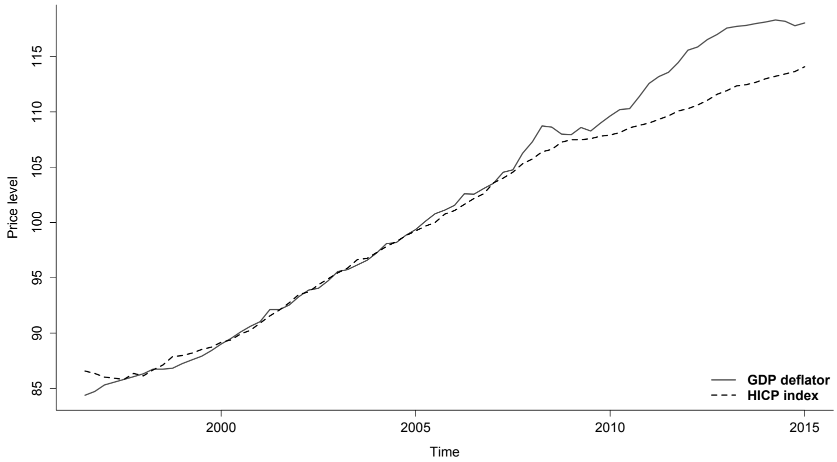

Figure 7.1: Development of price indexes for EU 18

the GDP deflator consumer index showcased almost constant growth. The development of these series was nevertheless very similar until the crisis where the difference started to grow. Currently, the deflator is around 4 percentage points higher in terms of growth since 2005, compared to the HICP index.

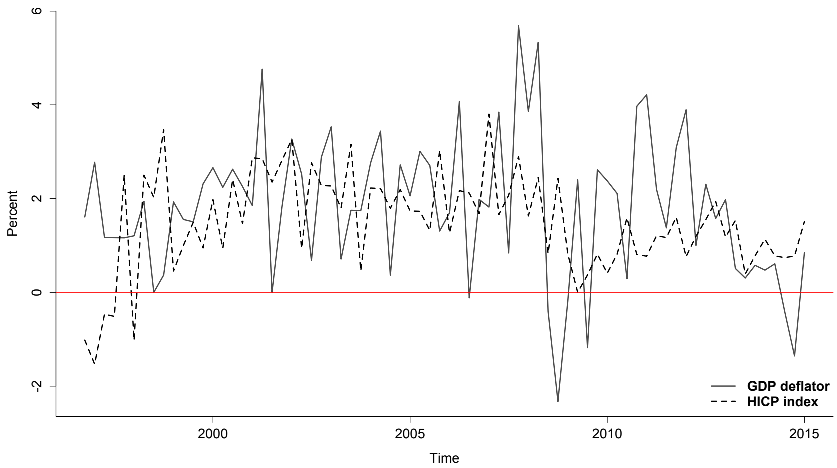

Figure 7.2: Development of price indexes for EU 18

Figure 7.2 shows the growth of the two series. The inflation was computed as a log difference of all two consecutive terms in each series which was furthermore multiplied by 100 to get the percentages. Moreover, the resulting series were multiplied by four to get the approximate yearly growth. The same approach can be found e.g. in Marcellino & Musso (2010). It can be seen in the graph that from the beginning of the century to the crisis, the behaviour of the two was relatively similar, however, since then the GDP deflator has started to become much more volatile relatively to the HICP index. Additionally, HICP index has never reached negative values since 1998 while the deflator got to the negative values three times during this period.

The different behavior of the two inflation proxy variables is also summarized in the following table:

	GDP deflator	HICP index
Min.	-2.33	-1.53
Median	1.83	1.54
Mean	1.81	1.49
Max	5.69	3.80
Std. dev.	1.51	1.06
ADF p-value	0.09	0.16
KPSS p-value	>0.1	0.05
Obs	74	74

Table 7.1: Summary statistics of inflation in EU 18

It is visible that all the statistics suggest higher dispersion of the GDP deflator. Its standard deviation is almost fifty percent bigger than the one of HICP index. Moreover, its minimum as well as the first quartile lies below that of the consumer index while the third quartile and maximum are higher. These findings suggest what can be partly seen in the previous figure that, in terms of growth, the HICP index is somehow enfolded by the deflator. In addition to this, the HICP inflation is on average approximately 0.3 percentage points lower. The fact that median is close to the mean for both variables suggests that their distributions exhibit low skewness.

Finally, Kwiatkowski-Phillips-Schmidt-Shin (KPSS) as well as Augmented Dickey-Fuller (ADF) tests suggested stationarity of HICP index and non-stationarity of GDP deflator. ADF test assumes unit root in its null hypothesis while KPSS sets it to be stationary under the null. This is the reason why a change in inflation, rather than inflation itself is forecasted in this work, as described in Chapter 6 and Marcellino & Musso (2010). After second differencing of series, the null of ADF test could be rejected at 2% level of significance and the null of KPSS was not rejected at 10% significance level which provides strong evidence in favour of stationarity.

Chapter 8

Dispersion decomposition of output gap estimates

The following analysis tackles the problem of output gap uncertainty along all dimensions explained in Table 2.1, except the parameter uncertainty which, on one hand, can cause bias in the model results, while on the other hand the goal of this thesis is to analyse the uncertainty regarding the real-time estimates to which parameter instability is of less significance, especially for the models used in this text.

The individual dimensions of uncertainty are, however, not evaluated separately in the following text. In other words, the evaluation strategies are not mutually exclusive. For example, the dispersion of real-time vintages captures the real-time corrections, end-sample statistical bias as well as the final estimation error. Part of this could be corrected by e.g. using quasi real-time data. On the other hand, multiple authors have tackled the problem of uncertainty from this perspective and the author believes that it would bring only marginal added value.

8.1 Uncertainty within models

Figure 8.1 captures the dispersion of real-time data vintages within models in the following way. The last observation of the vintage was deducted from its value 4, 8 and 12 vintages later and this process was rolled over the whole dataset. This resulted in obtaining three data vectors of 22 observations for each model. The box plots are constructed in a way that the bold black line corresponds to the mean of the vector, the box around corresponds to the range between the first and third quartile and the horizontal lines represent the constructed 95% confidence interval. There are several data points lying outside this range in the chart and a few more even further, especially for SVAR. The interval of -3 to 3 percentage points was chosen for the sake of clarity.

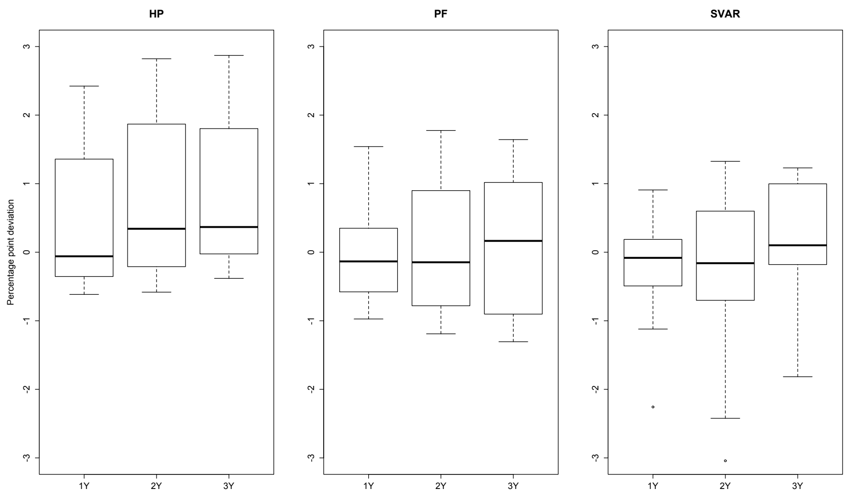

Figure 8.1: Development of price indexes for EU 18

As can be read from the previous graph, for all models in general, the highest difference in estimation occurs between the first and the ninth quarter. The change of uncertainty after the second year is significantly smaller when considering the confidence intervals as well as interquartile ranges. Hodrick-Prescott filter suffers from the largest dispersion during the first two years which is not surprising because of the end-sample biases of this model. The SVAR approach has, on one hand, the smallest interquartile range among the inquired models; while on the other hand, it seems to be producing some big outliers during the estimation process. The production function approach seems to be giving the relatively most consistent results among the three analysed shifts.

In addition to this graphical analysis, the author presents the statistics computed for each model in the following tables. Before proceeding to the analysis of the respective numbers, however, the author would like to introduce the last statistic, the same signs percentage. It is computed in the following fashion. Firstly, the direction of the output gap of a given observation is inquired which is then compared to the sign of the gap coming from vintages after 1, 2 and 3 years. The author believes that this kind of analysis may be of high importance for policy-makers because of the fact that the gap is unobservable. While there may be uncertainty regarding its exact value, the possible finding that at least the sign of the gap is consistent in time may help the policy-makers to assess the position of the economy, i.e. whether it is somehow overheating or underperforming. One last remark to the presented statistics is that one would a priory assume that all except the same signs percentage should grow with the growing shift in years while at the same time this growth should be declining. This comes from the belief that the more recent vintages include some additional information and therefore they change the behaviour of a whole series; however, their weight regarding older observations should decline in time.

	1Y	2Y	3Y
Min.	-0.62	-0.58	-0.38
Median	-0.06	0.34	0.37
Mean	0.44	0.77	0.89
Max	2.42	2.82	2.87
Std. dev.	1.01	1.16	1.12
Obs	22	22	22
Same signs pct.	83.36	81.82	81.82

Table 8.1: Statistics of real-time dispersion from HP gaps

The statistics of HP filter analysis suggest that the real-time vintages during were underestimating the gap on average by almost 1 percentage point. This bias was caused mainly around the crisis period, as can be seen in Figure D.1, which can be found in Appendix D. The chart depicts individual differences between vintages at given time. The median errors behaved similarly except for the 1 year shift where these were close to zero. The standard deviation seemed to stabilize at some level; same as maximum values while the minimum converged to zero.

In general, the HP filter seemed to produce relatively consistent results, with dispersion between the vintages lying in 0.5 percentage point interval around zero, only in the calmer periods before and after the crisis. To help the reader to assess the size of such mismeasurement, the half percentage point difference in output gap would currently make over 12 billion Euros which is a notable number for a policy-maker pursuing stabilization of economy. On the other hand, the statistic regarding percentage of the same signs gives more encouraging results in terms of output gap's qualitative analysis. In all vintages, over 80% of calculations resulted in the same direction. Moreover, Figure D.2 from Appendix D suggests that the most differences occurred during the crisis while the rest of the dataset was consistent in the qualitative behaviour of output gap.

	1Y	2Y	3Y
Min.	-0.97	-1.19	-1.31
Median	-0.13	-0.15	0.17
Mean	0.01	0.06	0.08
Max	1.54	1.78	1.64
Std. dev.	0.77	0.97	1.02
Obs	22	22	22
Same signs pct.	81.82	86.36	86.36

Table 8.2: Statistics of real-time dispersion from PF gaps

The production function statistics present slightly different story. While the dispersion between vintages was on average zero, the development was far from being steady, as can be seen in Figure D.3, Appendix D. The behaviour of these series before 2009 was similar to that of HP filter results, only on smaller scale, while the development after 2010 took larger magnitude resulting in over 1 percentage point overestimation of the gaps. The signs were also reversed around 2011 which was a period of another recession in euro area. Generally, the PF statistics suggest similar reliability of this method as that of Hodrick-Prescott filter.

	1Y	2Y	3Y
Min.	-6.75	-6.88	-6.70
Median	-0.08	-0.16	0.10
Mean	-0.33	-0.48	0.01
Max	3.38	1.33	1.23
Std. dev.	1.74	1.77	1.67
Obs	22	22	22
Same signs pct.	81.82	77.27	90.91

Table 8.3: Statistics of real-time dispersion from SVAR gaps

Completely different picture is given by Table 8.3 concerning the dispersion statistics of SVAR. The minimum of almost 7 percentage points was achieved by all 3 data vectors which truly is a huge overestimation error. However, this comes from the fact that one SVAR model was non-stationary, creating unrealistically large gap. The author chose to keep that model in the results as the goal of this thesis is to approach the real-time situations and these stability issues simply may occur. Interestingly, the SVAR model differences between vectors are extremely low after around half of 2009 regardless of the actual size of their individual differences. This may come from the fact that the vector autoregression model already 'learned' the data and was left to deal only with data revisions in the following vintages, omitting the statistical end-sample issues to which it should be relatively robust.

To sum up these findings, one has to be very careful when working with real-time data to estimate potential output and output gap. All models showed high values of dispersion when comparing real-time and ex-post data based gaps, especially during the Great Depression. The difference between the vintages was of much lower significance than actual position of the economy at the evaluated period. This basically means that data revision uncertainty is a smaller problem than statistical revision when evaluating output gap. The only model which was suggesting low vulnerability to real-time data was SVAR, nevertheless, it should be noted that this was the case only for period at the end of the sample, consisting of insufficient observations to do proper statistical test. On the other hand, the qualitative results suggest that all models have over 80% 'correct' directions of output gap evaluated in real-time. This information may help to give policy-makers some degree of assurance about where the economy is located at the time. The development of potential output estimates among vintages in time is depicted in charts of Appendix E.

8.2 Uncertainty between models

The uncertainty between the models is evaluated using the already presented statistic of the same signs. The directions of output gap for the whole dataset as well as for only the last 8 quarters of each vintage were inquired.

	HP vs. PF	HP vs. SVAR	PF vs. SVAR	ALL
Whole sample	91.90	84.15	80.00	77.77
Last 8Q	83.46	77.21	80.51	70.59
Last 1Q	70.59	82.35	70.59	61.76

Table 8.4: Percentage of same output gap directions

As could be expected, the real-time estimates have lower percentage than the whole sample statistics, except for the comparison of production function and structural VAR. Nevertheless, the real-time results of two model comparison are very similar to those obtained in the previous chapter, equaling around 80%. All models estimated the same direction of output gap almost 80% times for the whole sample, around 70% for the last 8 quarters and only approximately 60% for the final quarter of a vintage. While the whole sample results suggest that output gap modeling may be a nice tool for ex-post description of the economy's past behavior, the findings coming from the final quarter deteriorate this estimation to be of very little use when using it as a real-time measure of economic activity. Following these findings, the next chapter regarding inflation forecasting power of the models does not present surprising conclusions.

Chapter 9

Inflation forecasting power of output gap

9.1 HICP index

The results of corresponding tests are presented in Table 9.1 and Table 9.2. The null hypothesis that the difference between squared inflation change forecasts by autoregressive models and models incorporating output gap was tested against the one-sided alternative that this difference is positive i.e. that output gap models produce smaller squared errors. No single test resulted in rejecting the null hypothesis at any reasonable significance level.

	HP	PF	SVAR
1 quarter	0.13	0.15	0.76
4 quarters	0.99	1.00	0.94

Table 9.1: Student's p-values HICP index

The findings are actually in contrast to assumption that output gap should be significant predictor of inflation. Even more surprisingly, 4 quarter ahead forecast incorporating SVAR output gap is significantly, at ten percent level, worse than the autoregressive model and additionally, 4 period ahead forecast incorporating HP filter output gap is worse than autoregressive model when testing at 1% level of significance.

Another statistic - Root mean squared error (RMSE) is presented in Table 9.3. It is computed simply as square root of average squared error. Because of the fact, that benchmark AR model for SVAR methodology had 3 less observations than the benchmark model for PF and HP filter approaches, its statistic

	HP	PF	SVAR
1 quarter	0.46	0.51	0.76
4 quarters	0.99	0.99	0.92

Table 9.2: Rank test p-values HICP index

is presented separately as AR-s. It can be seen that autoregressive models produced very simple results in the 1 period ahead forecast as the models incorporating output gaps in terms of RMSE. For the four period ahead forecast, the AR models' RMSE was as much as one third lower than that of HP and PF modelel'.

	AR	AR-s	HP	PF	SVAR
1 quarter	0.18	0.18	0.15	0.16	0.20
4 quarters	0.22	0.22	0.32	0.31	0.24

Table 9.3: Root mean squared errors HICP index

9.2 GDP deflator

Utilization of GDP deflator brings similar results as that of HICP index. Both tests suggest AR models to be superior in predicting power in 4-period ahead forecast for HP filter and SVAR approach.

	HP	PF	SVAR
1 quarter	0.63	0.67	0.49
4 quarters	0.93	0.85	0.98

Table 9.4: Student's p-values GDP deflator

As was stated in Chapter 1, the prominent institutions utilize production function method most frequently in their estimates. Although this analysis did not prove usefulness of such estimates for forecasting inflation, the production function displayed relatively best results among models used which supports its favourability compared to other approaches.

	HP	PF	SVAR
1 quarter	0.53	0.59	0.42
4 quarters	0.96	0.79	0.97

Table 9.5: Rank test p-values GDP deflator

The RMSEs coming from the models using GDP deflator obtained higher values than their HICP counterparts which is, however, attributable to higher variance during the analysed period. Nevertheless, the results are again very similar to those obtained from analysis of HICP index.

	AR	AR-s	HP	PF	SVAR
1 quarter	2.20	2.22	2.30	2.33	2.19
4 quarters	2.01	2.01	2.11	2.07	2.12

Table 9.6: Root mean squared errors of GDP Deflator

Although using slightly different methodology, the results of this analysis are well in line with Marcellino & Musso (2010). The findings in this thesis suggest not only that output gap brings no additional value to inflation forecasting, but also that it can, in some cases, actually worsen the autoregressive models. Conclusively, among the economic models, the results in this thesis are consistent e.g. with the New-Keynesian Phillips Curve (NKPC) as specified by Roberts (1995) who suggests that the output gap has no forecasting power regarding inflation as the former leads, instead of lags, the latter.

SYNOPSIS

Discussion and summary

Chapter 10

Drawbacks

There are several ways how this thesis could be extended. The author tried to specify all the areas with space for improvement that he encountered during the work. Nevertheless, there is a good chance that some hidden flaws, not stated in this section, will be exposed by the reader.

First of all, there is one family of models quite frequently used in output gap estimation literature which was omitted from this work. Those are unobserved components models which utilize Kalman filter in order to detrend the output. One explanation for not utilizing this approach is that there are vast possibilities for specifications of these models. There actually is a possibility to, for example, nest the production function into unobserved component model. Another reason is that these models are much harder to implement and, in general, contain a larger set of parameters which may not be feasible with smaller data vintages used in this text because of parsimony rules. On the other hand, the proper specification of UC models may lead to significant inflation forecasting power of output gap as given e.g. by Proietti *et al.* (2007). Conclusively, the goal of this work is to stick to the most general specification of each type of models with Kalman filtering resulting to be beyond the scope of this text.

Secondly, when assessing the uncertainty of real-time data, this work does not utilize the concept of quasi real-time data. These are basically ex-post data from which shorter vintages are stripped and rolled over. By utilizing this concept, it would be possible to separate the data revisions uncertainty while statistical end-sample biases would be kept. On the other hand, as was stated by Camba-Mendez & Rodriguez-Palenzuela (2001), the data revisions themselves are of lower significance, with the main portion of uncertainty in real-time data being surrounding the statistical revision.

One other possible question may be raised regarding the methodology of inflation forecasting. At some point this work deflected from the method utilized by Marcellino & Musso (2010) who used the special MSE-t statistic developed directly for testing the real-time data. The author of this work believes that firstly, the construction of this statistic is beyond the scope of this text and secondly, the results of output gap's inflation forecasting power are so strong that even bias of medium magnitude in the statistics displayed in previous chapter would not change them.

Chapter 11

Conclusion

The thesis inquired the output gap development in the euro area, tackled the problems connected with estimation of this latent variable and evaluated its inflation forecasting power. This chapter will provide both qualitative and quantitative conclusions resulting from previous analysis. Generally, the findings regarding output gap evolution and inflation prediction are in line with previous research. Additionally, this work quantified the uncertainty surrounding output gap estimation in real-time along several dimensions.

The first major finding comes from the evaluation of output gap during the analysed period from 1998 to 2014. The hysteresis effect, observed e.g. by Ball (2014) on the study of OECD countries or Reifschneider *et al.* (2013) in the US, is most likely present in the euro area as well. The growth of potential output computed in this work was close to 2% per annum until 2007. Since then, the potential has been suffering from stagnation, averaging around zero yearly growth. This is in line with the current euro area research conducted by Anderton *et al.* (2014) who suggested that there may be medium term damage unless some structural actions are taken. Important note is that growth rates of potential output resulting from all models used in this text lay in some small range when inquiring the analysed period ex-post. These findings reinforce the idea that the Great Depression caused structural change in the economy which caused large damage to its potential, an assertion, which was proved by formal test of structural break. All models suggested that the effects of crisis were already present in late 2007 or early 2008. The small range of structural break identification by statistical analysis as well as low dispersion of potential output growth rates brings additional confidence into ex-post estimation of potential output.

Next part of the thesis was concerning the problems of real-time output gap computation. On one hand, ex-post evaluation of potential output may be an elegant tool for evaluation of economy's position. On the other hand, the question may be raised whether there is some usefulness in its real-time applications. The problem has been tackled before by e.g. Orphanides & van Norden (2002), Marcellino & Musso (2010) and great many other studies which mostly resulted in low reliability of real-time data estimation. This thesis inquired the development of various forms of uncertainty in time concluding that data revisions obtain smaller degree of uncertainty than statistical revision as the dispersion among estimates of individual vintages mostly depends on the time period when they are being evaluated. The range of dispersion during the analyzed period resulting from revision after 8 quarters was around 3 percentage points for Hodrick-Prescott filter and production function approach while SVAR method provided even more extreme values. These findings suggest high degree of uncertainty in the euro area's output gap given that each percentage difference in potential output evaluation corresponds to around 24 billion Euros.

Given the fact that real-time quantitative measures of output gap are surrounded by very high degree of uncertainty, the author of this work inquired whether at least its qualitative assessment in real time provides some consistent results in euro area. For example, at the end of the analysed period, all of the three models suggested different direction of output gap. While production function approach provided negative gap because of high level of unemployment, Hodrick-Prescott filter showed zero and SVAR method resulted in positive output gap. The level of consistency was checked within as well as between models using the percentage of same signs measure. Interesting finding, that among vintages within models, the percentage of same signs was mostly over 80% may bring some more confidence for policy-makers evaluating the gap in real-time as, although not being able to decide on the actual size of the gap, they may possibly at least rely on its direction most of the time.

These results, however, suggest only relative consistency of output gap signs within models. The consequences of sign analysis between models are again hinting low reliability of real-time data. While the directions of the signs between any two models are the same about 80% times in ex-post evaluation as well as in the evaluation of last 8 quarters, the comparison of all three models results in less than 80% and around 70% respectively. Additionally, when inquiring only the last quarter of a vintage, all models agree on the direction of the gap in only 60% of cases. For euro area policy-makers responsible for the period analysed in this work, this would mean that they would be exposed to model uncertainty for almost half of their decisions regarding output gap direction.

The last part of analysis was concerning the inflation prediction power of output gap. As could be expected from previous findings regarding real-time data vintages, the variable did not result to be a significant predictor of inflation change in any specification. The results are very similar to those by Marcellino & Musso (2010). There are at least two consequences of these findings. Firstly, implicit relationship between output gap and inflation can be found in e.g. Taylor rule or by combination of Okun's law and Phillips curve. Analyses, such as this one, suggest that despite that these economic phenomena may be observed when output gap and inflation are evaluated at the same time period, however, when output gap leads inflation; the models do not work, as suggested by e.g. Roberts (1995) in his NKPC. Secondly, the contribution is made to the inflation flattening discussion, which was opened e.g. by IMF (2013) as autoregressive models performed the inflation forecast significantly better than models incorporating output gap.

Generally, output gap seems to be untrustworthy measure of economic activity in euro area when estimated in real-time. Additionally, it does not bring any added value to autoregressive models of inflation forecasting. On the other hand, potential output analyzed ex-post gives consistent results within as well as among the models and can help describe the past behavior of the economy.

Bibliography

Aaronson, S., T. Cajner, B. Fallick, F. Galbis-Reig, C. Smith, & W. Warscher (2014): "Labor Force Participation: Recent Developments and Future Prospects." *Federal reserve bank of cleveland*, Federal Reserve Bank of Cleveland.

Adjemian, S., M. D. Paries, & S. Moyen (2007): "Optimal Monetary Policy in an Estimated DSGE for the Euro Area." *ECB Working Paper Series 803*, European Central Bank.

Anderton, R., T. Aranki, A. Dieppe, C. Elding, S. Haroutunian, P. Jacquinot, V. Jarvis, V. Labhard, D. Rusinova, & B. Szorfi (2014): "Potential Output from a Euro Area Perspective.Potential Output from a Euro Area Perspective." *ECB Occasional paper series 156*, European Central Bank.

Atkeson, A. & L. E. Ohanian (2001): "Are Phillips Curves Useful for Forecasting Inflation?" *Federal Reserve Bank of Minneapolis Quarterly Review* **25(1)**: pp. 2–11.

Ball, L. M. (2014): "Long-Term Damage from the Great Recession in OECD Countries." *NBER Working Paper 20185*, National Bureau of Economic Research.

Baxter, M. & R. G. King (1999): "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series." *Review of Economics and Statistics* **81**: pp. 575–593.

Bayoumi, T. & B. Eichengreen (1992): "Is There a Conflict between EC Enlargement and European Monetary Unification?" *NBER Working Paper 3950*, National Bureau of Economic Research.

Blanchard, O. J. & D. Quah (1989): "The Dynamic Effects of Aggregate Demand and Aggregate Supply." *The American Economic Review* **79(4)**: pp. 655–673.

Bouis, R., B. Cournede, & A. K. Christensen (2012): "Implications of Output Gap Uncertainty in Times of Crisis." *OECD ECO Working Papers 977*, Organization for Economic Co-operation and Development.

Camba-Mendez, G. & D. Rodriguez-Palenzuela (2001): "Assessment Criteria for Output Gap Estimates." *ECB Working Paper 54*, European Central Bank.

Cass, D. (1965): "Optimum Growth in an Aggregative Model of Capital Accumulation." *Review of Economic Studies* **32(3)**: pp. 233–240.

Cerra, V. & S. C. Saxena (2000): "Alternative Methods of Estimating Potential Output and the Output Gap: An Application to Sweden." *IMF Working paper 59*, International Monetary Fund.

Clarida, R., J. Gali, & M. Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Theory and Some Evidence." *Quarterly Journal of Economics* **115(1)**: pp. 147–180.

Claus, I. (1999): "Estimating Potential Output for New Zealand: A Structural VAR Approach." *Discussion Paper Series 3*, Reserve Bank of New Zealand.

Derbyshire, J., B. Gardiner, & S. Waights (2010): "Estimating the Capital Stock for the NUTS 2 Regions of the EU-27." *EU Regional Policy Working Paper 01/2011*, European Commission.

Fernald, J. (2014): "Productivity and Potential Output Before, During, and After the Great Recession." *NBER Macroeconomics Annual*, National Bureau of Economic Research.

Fisher, I. (1896): "Appreciation and Interest." *American Economic Association* **XI(4)**: pp. 331–442.

Giorno, C., P. Richardson, D. Roseveare, & P. van den Noord (1995): "Potential Output, Output Gaps and Structural Budget Balances." *OECD Economic Studies 24*, Organization for Economic Co-operation and Development.

Gordon, R. J. (2013): "The Phillips Curve is Alive and Well: Inflation and the NAIRU During the Slow Recovery." *NBER Working Paper 19390*, National Bureau of Economic Research.

Hall, R. E. (2014): "Quantifying the Lasting Harm to the U.S. Economy from the Financial Crisis." *NBER Working Paper 20183*, National Bureau of Economic Research.

Havik, K., K. McMorrow, F. Orlandi, C. Planas, R. Raciborski, W. Roger, A. Rossi, A. Thum-Thysen, & V. Vandermeulen (2014): "The Production Function Methodology for Calculating Potential Growth Rates & Output Gaps." *European Economy Economic Papers 535*, European Commission.

Hodrick, R. J. & E. C. Prescott (1997): "Postwar U.S. Business Cycles: An Empirical Investigation." *Journal of Money, Credit and Banking* **29**: pp. 1–16.

IMF (2009a): "The Globalization of Labor." *World economic outlook*, International Monetary Fund.

IMF (2009b): "What's the Damage? Medium-Term Output Dynamics after Financial Crises." *World economic outlook*, International Monetary Fund.

IMF (2013): "The Dog That Didn't Bark: Has Inflation Been Muzzled or Was It Just Sleeping?" *World economic outlook*, International Monetary Fund.

Kaiser, R. & A. Maravall (1999): "Estimation of the business cycle: A Modified Hodrick-Prescott Filter." *Spanish Economic Review* **1**: pp. 175–206.

Kiley, M. T. (2010): "Output Gaps." *FEDS Workng Papers 27*, Federal Reserve Board.

King, R. G., C. I. Plosser, J. H. Stock, & M. W. Watson (1991): "Stochastic Trends and Economic Fluctuations." *The American Economic Review* **81**: pp. 819–840.

Koopmans, T. C. (1963): "On the Concept of Optimal Economic Growth." *Cowles Foundation Discussion Papers 163*, Yale University.

LUCAS, R. E. J. & L. A. RAPPING (1969): "Price Expectations and the Phillips Curve." *The American Economic Review* **59(3)**: pp. 342–350.

MANKIW, G. N., D. ROMER, & D. N. WEIL (1992): "A Contribution to the Empirics of Economic Growth." *The Quarterly Journal of Economics* **107(2)**: pp. 407–437.

MARCELLINO, M. & A. MUSSO (2010): "Real Time Estimates of the Euro Area Output Gap: Reliability and Forecasting Performance." *ECB Working paper series 1157*, European Central Bank.

OECD (2001): "Measuring Capital: Measurement of Capital Stocks, Consumption of Fixed Capital and Capital Services." *Oecd manual*, Organization for Economic Co-operation and Development, Paris.

OECD (2009): "Measuring Capital." *Oecd manual*, Organization for Economic Co-operation and Development, Paris. Second Edition.

OINONEN, S. & M. PALOVIITA (2014): "Updating the Euro Area Phillips Curve: The Slope has Increased." *Research Discussion Papers 31*, Bank of Finland.

OKUN, A. M. (1962): "Potential GNP: Its Measurement and Significance." *Cowles Foundation Paper 190*, Yale University.

ORPHANIDES, A. (2003): "The Quest for Prosperity without Inflation." *Journal of Monetary Economics* **50(3)**: pp. 633–663.

ORPHANIDES, A. & S. VAN NORDEN (2002): "The Unreliability of Output Gap Estimates in Real Time." *Review of Economics and Statistics* **84(4)**: pp. 569–583.

PHILLIPS, A. W. (1958): "The Relationship between Unemployment and the Rate of Change of Money Wages in the United Kingdom 1861-1957." *Economica* **25(100)**: pp. 283–299.

PROIETTI, T., A. MUSSO, & T. WESTERMANN (2007): "Estimating Potential Output and the Output Gap for the Euro Area: a Model-Based Production Function Approach." *Empirical Economics* **33**: pp. 85–113.

QUINT, D. & P. RABANAL (2013): "Monetary and Macroprudential Policy in an Estimated DSGE Model of the Euro Area." *IMF Working Paper 209*, International Monetary Fund.

Ramsey, F. P. (1928): "Mathematical Theory of Saving." *Economic Journal* **38(152)**: pp. 543–559.

Reifschneider, D., W. L. Wascher, & D. Wilcox (2013): "Aggregate Supply in the United States: Recent Developments and Implications for the Conduct of Monetary Policy." *FEDS Working Papers 77*, Federal Reserve Board.

Roberts, J. M. (1995): "Keynesian Economics and the Phillips Curve." *Journal of Money, Credit and Banking* **27(4)**: pp. 975–984.

Roberts, J. M. (2006): "Monetary Policy and Inflation Dynamics." *International Journal of Central Banking* **2**: pp. 193–230.

Romer, P. M. (1986): "Increasing Returns and Long-Run Growth." *Journal of Political Economy* **94(5)**: pp. 1002–1037.

Sargent, T. J. (1981): "The Ends of Four Big Inflations." *Working Paper 158*, Federal Reserve Bank of Minneapolis and University of Minnesota.

Sims, C. (1980): "Macroeconomics and Reality." *Econometrica* **48(1)**: pp. 1–48.

Smets, F. (1998): "Gap Uncertainty: Does It Matter for the Taylor Rule?" *BIS Working Papers 60*, Bank for International Settlements.

Sollow, R. M. (1956): "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* **70(1)**: pp. 65–94.

Stock, J. H. & M. W. Watson (1999): "Business Cycle Fluctuations in U.S. Macroeconomic Time Series." *NBER Working Paper Series 6528*, National Bureau of Economic Research, Cambridge.

Stock, J. H. & M. W. Watson (2008): "Phillips Curve Inflation Forecasts." *Technical report*, Federal Reserve Bank of Boston.

Swan, T. W. (1956): "Economic GGrowth and Capital Accumulation." *Economic Record* **32(2)**: pp. 334–361.

Taylor, J. B. (1993): "Discretion Versus Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy* **39(1)**: pp. 195–214.

Torres, R., P. Jarrett, & W. Suyker (1989): "Modelling Business Sector Supply for the Smaller OECD Countries." *Economics and Statistics Department WP 71*, Organization for Economic Co-operation and Development.

Torres, R. & J. P. Martin (1989): "Potential Output in the Seven Major OECD Countries." *OECD Economic Studies 14*, Organization for Economic Co-operation and Development.

Wold, H. (1939): "A Study in the Analysis of Stationary Time Series." *Journal of the Royal Statistical Society* **102(2)**: pp. 295–298.

LIST OF APPENDICES

Appendix A

Importance of output gap and potential output

There is a vast number of macroeconomic laws, especially regarding business cycles and their relation with inflation and unemployment, incorporating the output gap variables. Based on the author's best knowledge this text summarizes the most important models for the work.

A.1 Sollow-Swan model

Sollow-Swan is perhaps the most widely known macroeconomic model. It can be viewed as the first brick in building the theory of growth. Independently introduced by Sollow (1956) and Swan (1956), the model is described as follows:

$$Y(t) = K(t)^{\alpha} \left(A(t)L(t)\right)^{1-\alpha}, \tag{A.1}$$

where

t = time
$Y(\cdot)$ = production of the economy
$K(\cdot)$ = capital input
$L(\cdot)$ = labour input
A = labour-augmenting technology
α = elasticity of output with respect to capital

The parameter α is assumed to lie strictly between zero and 1. The levels of labour and technology are subject to exogenous growth at rate n and g respectively, as given in the following equation.

$$L(t) = L(0)e^{nt} \tag{A.2}$$

$$A(t) = A(0)e^{gt} \tag{A.3}$$

Moreover, capital depreciates at rate δ. This is being offset by savings of households who can exchange their part current consumption, s such that $0 \leq s \leq 1$, for next period's capital, as given here:

$$\dot{K}(t) = sY(t) - \delta K(t) \tag{A.4}$$

$$\dot{K}(t) = \frac{dK(t)}{dt} \tag{A.5}$$

The model assumptions are that the closed economy has diminishing returns to labour and capital and constant returns to scale. The saving ratio and technological progress are assumed to be constant and production factors are substitutable. Following that the model is classified as neo-classical, the prices and wages are assumed to be fully flexible, the factors of production are paid according to their marginal productivities, labour and capital are moreover fully employed.

The model works such that households choose optimal savings rate which maximizes their consumption in time. One of the real-life implications of the model is that the government or central bank can, through various policies, affect saving rates of the households. The shock to saving rate hits other variables of the model as well, perhaps most interestingly, the growth rate of output per unit of capital firstly shifts in the direction of the shock and then slides back to its balanced growth path resulting in conclusion that long-term growth cannot be affected and shocks to saving rate only cause short-run fluctuations.

There are numerous extensions to this model, starting with endogenous savings theory introduced by Ramsey (1928), Cass (1965) and Koopmans (1963), AK models, e.g. by Romer (1986), dealing with absence of diminishing returns to capital or models with the addition of human capital, e.g. by Mankiw *et al.* (1992). The Sollow-Swan model is important to this work because it provides economic background to the Production Function modelling of potential output. The following phenomenon provides also strong empirical finding supporting the essentiality of output gap estimation for the description of economy's behaviour.

A.2 Okun's law

Okun's law is one of the most famous economical relationships important for this work. Introduced by Okun (1962), this law captures the reverse co-movement of unemployment and output gaps. The equation is given as:

$$\frac{(Y - \bar{Y})}{\bar{Y}} = c(u - \bar{u}) \tag{A.6}$$

The fact that the empirical correlation of these variables is very strong was also shown by e.g. Stock & Watson (1999). The implicit conclusion that large portion of economy's cyclical behaviour can be explained by fluctuations in the labour market provided a lot of motivation for economic research to concentrate on labour economics. Two of three models used to estimate output gap in this thesis incorporate the factor of unemployment, the use of Production Function and Blanchard & Quah (1989) type SVAR model in this work is therefore reinforced by the strong economical relationship given by Okun's law. The Hodrick-Prescott filter is the third method, chosen to represent univariate models in order to compare them to multivariate models. The results then may show the usefulness of adding unemployment variables into potential output estimates. On the other hand, Okun's law can only be interpreted as an empirical finding without sufficient background in the economic theory.

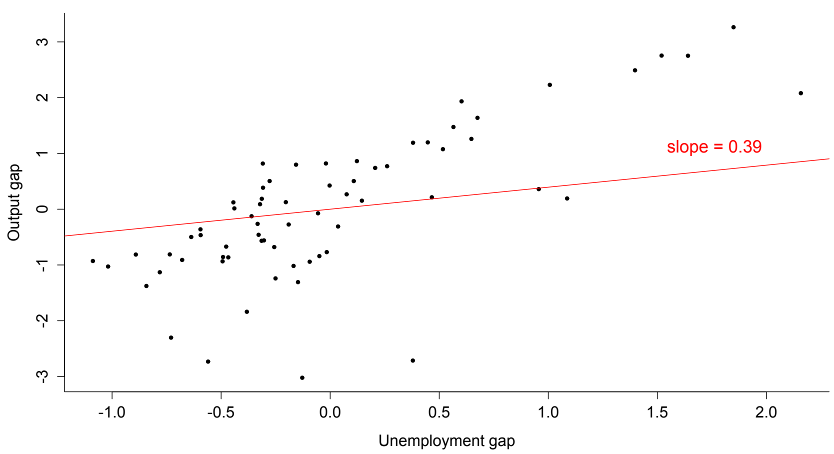

Figure A.1: Okun's law in euro area

Figure A.1 represents the Okun's relationship based on the data used in this work. It shows the co-movement of output gap obtained from Hodrick-Prescott filter evaluated at the December 2014 vintage and unemployment gap, resulting from HP filtering of the series as well. Quarterly data from OECD and EC were used since 1998 and the values on axes represent the percentage deviation from trend. The correlation between the two is certain. Additionally, when unemployment gap rises by 1 percentage point, output gap grows, on average, by around 0.4 percentage points. The linear regression moreover suggests very high statistical significance of this relationship with p-value being much lower than 0.001.

A.3 Taylor rule

Another famous and widely known economic law is most certainly the notorious Taylor rule. John B. Taylor (1993) specified the equation:

$$i_t = \pi_t + \alpha_\pi(\pi_t - \bar{\pi}_t) + \alpha_y(y_t - \bar{y}_t) \tag{A.7}$$

where

i = sovereign funds rate

π = inflation rate of GDP deflator

$\bar{\pi}$ = inflation target of central bank

y = real GDP

$\bar{y}$ = potential output

α_π = adjustment sensitivity to inflation

α_y = adjustment sensitivity to output

Basically, this is extension of Fisher equation, introduced by Fisher (1896), where nominal interest rates equal to the sum of inflation and real interest rates. The Taylor rule builds on it by adding the inflation and output gaps. The inflation gap here is defined as the difference between actual inflation and its target while output gap definition is consistent with the one given earlier in the text. This formulation provides nice implications regarding the monetary policy.

In his work, Taylor (1993) specified the exact numbers for coefficients in the model which well described the behaviour of US nominal interest rates for the period between 1987 and 1992. He moreover argued that when actions of the central banks are based on rules, like Taylor rule, instead of discretionary decrees, the credibility of these central banks in the public should be fostered because of lower uncertainty surrounding their future moves.

On one hand, many economists agree that the Taylor rule describes the movement of sovereign rates, especially in the USA and other developed countries, this is nicely summarized by Clarida *et al.* (2000). On the other hand, Orphanides (2003) argues that when real-time data vintages are used, the results' power drop significantly. This result is important for the following analysis of the thesis. In the original paper, Taylor proposes relatively high values for the gaps' coefficients, both suggested to be 0.5 which would result in somewhat aggressive monetary policy actions for adjustment of the economy.

The problem in this method may occur in the duality of the central bank goals. This occurs for example during the periods of stagflation with high inflation and low output. Interestingly, during the period of the Great Depression and even after the turmoil, the Taylor rule has suggested negative funds rate, either emphasizing the magnitude of the slack or disregarding its functionality. The Taylor rule is the first macroeconomic phenomenon which underlines the importance of inflation analysis in this paper.

A.4 Phillip's curve

There are other inflation - output gap economic relations present in the theoretical literature. Firstly, the Phillips curve is introduced here. There are many specifications of the curve taking different forms. The two traditional benchmarks are classical, based on Phillips (1958) and its New Keynesian version, derived by Roberts (1995). From the classical specification Gordon triangle model described later can be derived as well as the so-called new classical form based on the aggregate supply function introduced by Lucas & Rapping (1969).

The first, classical, Phillips curve described simply the wage acceleration as a function of unemployment and the trend growth rate of money wages. This can be extended to include also inflation expectations and NAWRU. This extension by Lucas & Rapping (1969) brings the information regarding the aggregate supply into the model which is of great importance for this work, as the unemployment gap implicitly includes the output gap, creating its link to inflation.

Possibly even more interesting results come from NKPC, introduced by Roberts (1995).The main implication of this model is that the inflation leads measures of the output gap. This would mean that inflation forecasting using the gap was basically useless.

It should be nevertheless noted that the Phillips curve is only an empirical finding and since this relationship was broken in 1970s, it has been criticised by great many authors, including the Nobel Prize winners, e.g. recent winner Sargent (1981). Some authors, as Atkeson & Ohanian (2001) or Stock & Watson (2008) even argue, that inflation is well described and forecasted by random walk models thus boosting the uncertainty regarding the predictions. On the other hand, the topic of Phillips curve is still kept alive, e.g. by Oinonen & Paloviita (2014) or Gordon (2013) who proposed that his alternative version described in the following chapter fits the data well. Moreover, he argued against the flattening of Phillips curve, proposed e.g. by Roberts (2006).

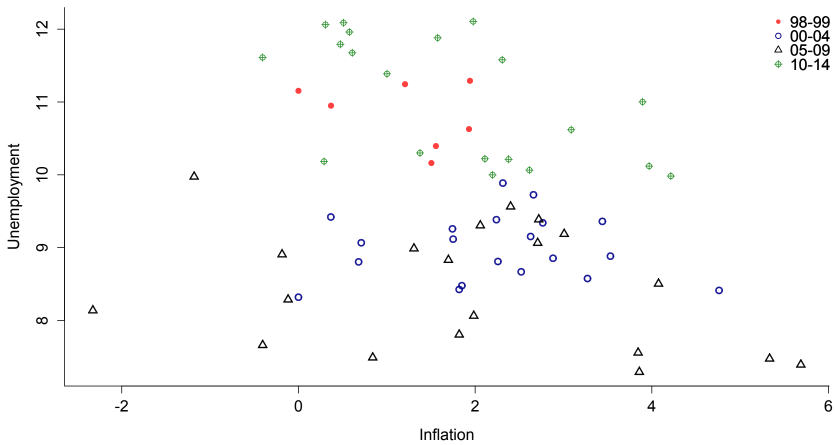

Figure A.2: Phillips curve in euro area

Figure A.2 shows euro area's Phillip's curve relationship since 1998 until 2014 on EC data. While throughout the most of this period, the relationship was really flat, there are some sings of negative relationship in the last five years. Based on the previous analysis of inflation - output gap relationship, the author a priory expected that re-estimating the models by Marcellino & Musso (2010) could provide fresh and possibly more positive results regarding the output gap's inflation forecasting performance.

A.5 Triangle model of inflation

Robert J. Gordon used the classical Phillips curve to describe the triangle model which understands three sources of inflation. According to this model, there can be demand-pull, cost-push or built-in inflation.

Demand-pull inflation arises when there are shifts to the aggregate demand. These shifts may be caused by e.g. increases in households' disposable income, government spending, interest rates etc. as given by the Keynesian economic stream. This type of inflation utilizes the Phillips curve relationship in the following way: as aggregate demand rises, more labour is required, which, according to the Phillips curve relationship, should cause the inflation to rise as well.

Cost-push inflation then arises in times of supply shocks to the economy. One notorious example is the oil crisis of the 1970s, where the cost of petroleum jumped up, causing huge problems in the supply chains of firms. This type of inflation causes Phillips curve to shift, in contrast with demand-pull inflation which moves along the curve.

The last, built-in inflation, according to its name, is based on its past behaviour. Among its determinants, inflationary expectations or price/wage spiral play an important role. It is backed by empirical evidence that inflation is highly persistent in nature, the evidence shown e.g. in IMF (2013).

The criticism of the Triangle model as well as Phillip's curves comes mostly from the monetarist economists. For example, Milton Friedman's famous statement about inflation being strictly monetary phenomenon which can be only caused by over creation of money in excess of demand is in the contrary with this theory.

Appendix B

SVAR methodology

Blanchard & Quah (1989) proposed a structural VAR method with long-run restrictions on effects of unemployment u on output y. The stationary version of output in log-differences is utilized in this approach, noted as Δy. The author of this thesis believes that description of SVAR methodology by Claus (1999) is especially well written as it is deep yet comprehensive summary of the approach. In his paper, Claus (1999) nevertheless used an extension of basic Blanchard-Quah type SVAR model and therefore, its re-transformed version is depicted here.

Firstly, it is important to realize, that according to theorem introduced by Wold (1939), any stationary process can be re-written in the moving average representation. The sequences of $\{\Delta y_t\}$ and $\{u_t\}$ can therefore be expressed as a linear combination of their past and current shocks in the traditional VAR model specification:

$$\Delta y_t = \sum_{k=0}^{\infty} s_{11}(k) v_{1t-k} + \sum_{k=0}^{\infty} s_{12}(k) v_{2t-k} \tag{B.1}$$

$$u_t = \sum_{k=0}^{\infty} s_{21}(k) v_{1t-k} + \sum_{k=0}^{\infty} s_{22}(k) v_{2t-k} \tag{B.2}$$

which, in matrix notation can be written as:

$$\begin{bmatrix} \Delta y_t \\ u_t \end{bmatrix} = \begin{bmatrix} S_{11}(L) & S_{12}(L) \\ S_{21}(L) & S_{22}(L) \end{bmatrix} \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix} \tag{B.3}$$

or

$$x_t = S(L) v_t \tag{B.4}$$

The shocks υ_t are normalized in the following way:

$$E(\upsilon_t \upsilon_t) = \begin{bmatrix} var(\upsilon_{1t}) & cov(\upsilon_{1t}, \upsilon_{2t}) \\ cov(\upsilon_{2t}, \upsilon_{1t}) & var(\upsilon_{2t}) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \tag{B.5}$$

υ_{1t} stands for shock to aggregate supply in time t while υ_{2t} represents an aggregate demand shock. What distinguishes SVAR from VAR models is that they restrict these shocks in a way which is convenient for the researcher, e.g. from the economic perspective. Blanchard-Quah type SVAR assumes that there is no long-term effect of aggregate demand shocks on aggregate supply, i.e. as given by Equation B.6, the sum of effects of υ_{2t} on Δy_t in time is restricted to zero.

$$\sum_{k=0}^{\infty} s_{12}(k) \upsilon_{2t-k} = 0 \tag{B.6}$$

The structural shocks υ_t are nevertheless latent variables. To obtain these one must first estimate the unrestricted VAR model of the form:

$$\begin{bmatrix} \Delta y_t \\ u_t \end{bmatrix} = \begin{bmatrix} \Phi_{11}(L) & \Phi_{12}(L) \\ \Phi_{21}(L) & \Phi_{22}(L) \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ u_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix} \tag{B.7}$$

or

$$x_t = \Phi(L) x_{t-1} + \epsilon_t \tag{B.8}$$

The same matrix of regressors are shared by all equations and therefore, after inquiring the optimal number of lags, reduced-form model can be estimated by application of Ordinary Least Squares (OLS) to each equation in B.7 separately. After obtaining the residuals, the model can be inverted to the Wold moving average representation, given by:

$$\begin{bmatrix} \Delta y_t \\ u_t \end{bmatrix} = \begin{bmatrix} C_{11}(L) & C_{12}(L) \\ C_{21}(L) & C_{22}(L) \end{bmatrix} \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix} \tag{B.9}$$

or equivalently

$$x_t = C(L) \epsilon_t \tag{B.10}$$

where

$$C(L) = (I - \Phi(L))^{-1}$$

Assumption is raised on that the some form of linear combination of structural disturbances υ_t generates the innovations in ϵ_t, as shown in the following equation.

$$\begin{bmatrix}\epsilon_{1t}\\ \epsilon_{2t}\end{bmatrix} = \begin{bmatrix}s_{11}(0) & s_{12}(0)\\ s_{21}(0) & s_{22}(0)\end{bmatrix}\begin{bmatrix}\upsilon_{1t}\\ \upsilon_{2t}\end{bmatrix} \tag{B.11}$$

or

$$\epsilon_t = S(0)\upsilon_t \tag{B.12}$$

It follows from Equation B.12 that

$$E(\epsilon_t\epsilon_t') = S(0)E(\upsilon_t\upsilon_t')S'(0) = \Sigma \tag{B.13}$$

Given that variance and covariance of reduced-form residuals is observed, this allows the researcher to find 3 equations of 4 unknowns resulting from Equation B.14.

$$\begin{aligned} E(\epsilon_t\epsilon_t) &= \begin{bmatrix} var(\epsilon_{1t}) & cov(\epsilon_{1t},\epsilon_{2t})\\ cov(\epsilon_{2t},\epsilon_{1t}) & var(\epsilon_{2t})\end{bmatrix}\\ \Sigma &= \begin{bmatrix} s_{11}^2(0)+s_{12}^2(0) & s_{11}^2(0)s_{21}^2(0)+s_{12}^2(0)s_{22}^2(0)\\ s_{11}^2(0)s_{21}^2(0)+s_{12}^2(0)s_{22}^2(0) & s_{21}^2(0)+s_{22}^2(0)\end{bmatrix} \end{aligned} \tag{B.14}$$

The last equation can be obtained from combining Equation B.3, Equation B.6, Equation B.9 and Equation B.11 from which one can get

$$\begin{bmatrix}S_{11}(L) & 0\\ S_{21}(L) & S_{22}(L)\end{bmatrix} = \begin{bmatrix}C_{11}(L) & C_{12}(L)\\ C_{21}(L) & C_{22}(L)\end{bmatrix}\begin{bmatrix}S_{11}(0) & S_{12}(0)\\ S_{21}(0) & S_{22}(0)\end{bmatrix} \tag{B.15}$$

The resulting four equations of four unknowns are then

$$\begin{aligned} &1.\ \ var(\epsilon_{1t}) &&= s_{11}^2(0)+s_{12}^2(0)\\ &2.\ \ var(\epsilon_{2t}) &&= s_{21}^2(0)+s_{22}^2(0)\\ &3.\ \ cov(\epsilon_{1t},\epsilon_{2t}) &&= s_{11}^2(0)s_{21}^2(0)+s_{12}^2(0)s_{22}^2(0)\\ &4.\ \ 0 &&= s_{12}(0)C_{11}(L)+s_{22}(0)C_{12}(L) \end{aligned}$$

The output series estimated by Blanchard & Quah (1989) type of SVAR can be decomposed into two components. The potential growth is constructed from all past and present structural shocks to aggregate supply while output gap is obtained by summing all demand side structural shocks in time, as shown in Equation B.16.

$$\begin{aligned} \Delta y_t^p &= S_{11}(L)\upsilon_{1t} \\ gap_t &= S_{12}(L)\upsilon_{2t} \\ \Delta y_t &= \Delta y_t^p + gap_t \end{aligned} \tag{B.16}$$

What is left to mention is the way optimal number of lags was chosen in this thesis. To determine this, Akaike and Schwarz criterions were utilized. The search was nevertheless restricted to maximum of two lags given the sample length which resulted in assigning $VAR(2)$ model for every utilized vintage.

Appendix C

Method for identification of structural break

The methodology regarding determination of structural break occurrence was constructed on the following model:

$$y_t = \alpha + \beta * break + \gamma * F_t\left(1, \ldots, T\right) + \delta * break * F_t\left(t+1, \ldots, T\right), \quad \text{(C.1)}$$

where

$t \in (1, T)$.
y is the tested series.
α is intercept of model which only includes whole sample trend.
$F(\cdot)$ is a function which computes the trend, *slope*, of time specified in the argument with resulting vector evaluated as $t * slope$.
break is a dummy variable obtaining zeros until time $t-1$ and ones afterwards.

The significance of coefficient δ in Equation C.1 resulting from t-test then demonstrates whether a change in the relationship of the time series at the chosen date was statistically important. The author rolled this model throughout the whole sample and the date with the smallest p-value was chosen to be the time when structural break occurred as it provides the most statistical evidence of such event. Additional confidence in this approach is given by the fact that after performing the tests, p-values were strictly convex in time - falling from the beginning of the sample until reaching the bottom value to rise back towards the end of the sample. This was probably given by the fact that there was likely only one structural break present in the euro area potential output as well as that no shift in the series occurred.

Appendix D

Dispersion of real-time vintages in time

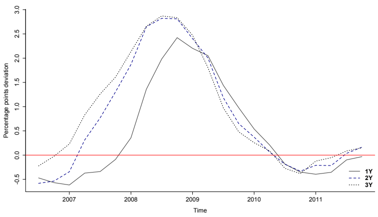

Figure D.1: HP filter dispersion

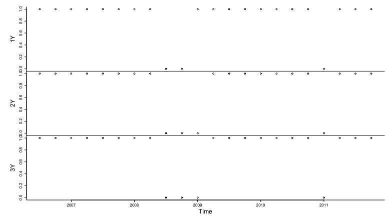

Figure D.2: HP filter same sign evaluation

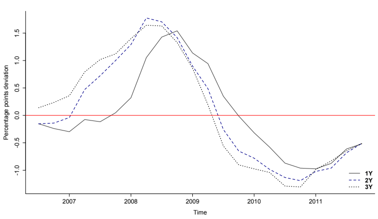

Figure D.3: PF dispersion

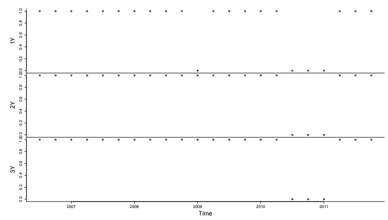

Figure D.4: PF same sign evaluation

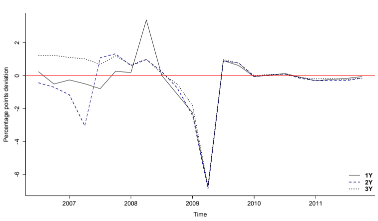

Figure D.5: SVAR dispersion

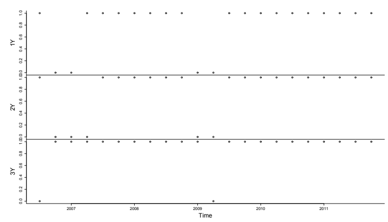

Figure D.6: SVAR same sign evaluation

Appendix E

Development of potential output estimates

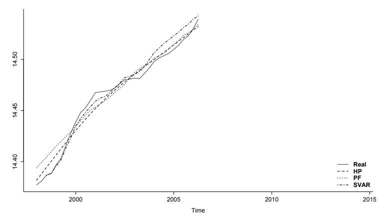

Figure E.1: September 2006 vintage

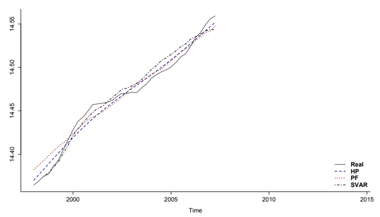

Figure E.2: September 2007 vintage

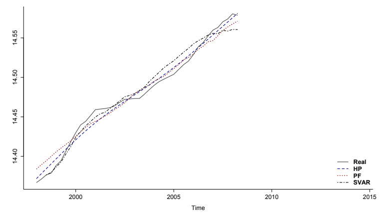

Figure E.3: September 2008 vintage

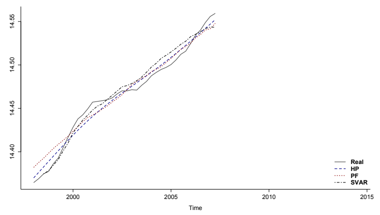

Figure E.4: September 2009 vintage

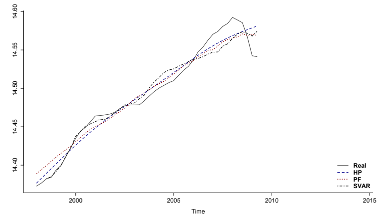

Figure E.5: September 2010 vintage

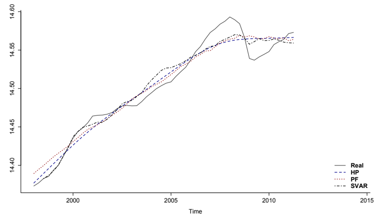

Figure E.6: September 2011 vintage

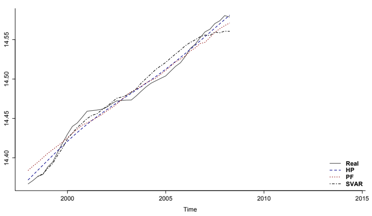

Figure E.7: September 2012 vintage

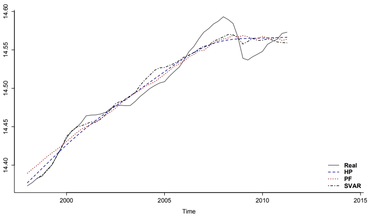

Figure E.8: September 2013 vintage

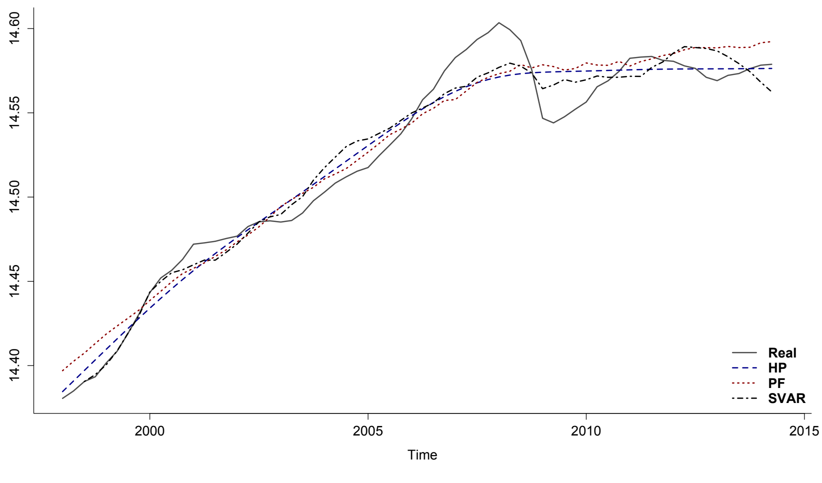

Figure E.9: September 2014 vintage